NOTE TO USERS

This reproduction is the best copy available.

UMI®

uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

Angela Yeung
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Microbiology and Immunology)
GRADE / DEGREE

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

Signature-tagged Mutagenesis in *Haemophilus ducreyi*
TITRE DE LA THÈSE / TITLE OF THESIS

C. Lee
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

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

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

B. Cameron

M. Lin

E. Hansen

T. Mah

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

Signature-tagged Mutagenesis in *Haemophilus ducreyi*

Angela Yeung

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirements for the PhD degree in Microbiology and Immunology

Department of Biochemistry, Microbiology and Immunology
Faculty of Medicine
University of Ottawa

© Angela Yeung, Ottawa, Canada, 2009

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-59492-6
Our file Notre référence
ISBN: 978-0-494-59492-6

NOTICE:

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

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

AVIS:

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

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

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

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

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

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

Canada

ABSTRACT

Signature-tagged mutagenesis (STM) is an *in vivo* mutation-based genetic screen that allows for the identification of virulence genes. A working model of STM in the temperature-dependent rabbit model (TDRM) of *Haemophilus ducreyi* has been established. A unique signature-tag (ST) was inserted upstream of the *aphA-3* gene in the transposon *Tn*1545-Δ3. Eleven different ST were used, each consisting of 21 unique nucleotides that were designed for PCR. A library of ST transposon *H. ducreyi* 35000 mutants was generated. A STM pool consisted of 10 unique ST-mutants that were inoculated intraepithelially into the TDRM. Ulcer development and histopathology of the lesions produced by the STM mutants inoculated at 10^5 CFU showed the same morphology as those produced by wild-type inoculated at 10^5 CFU. Ulcers that developed by day 4 were excised and recovered mutants in the output pool were screened by PCR. A previously characterized attenuated hemoglobin (HgbA) mutant was ST to serve as a positive control and was not recovered in the output pool. A total of 176 STM mutants were screened and 26 attenuated *H. ducreyi* mutants were recovered.

The 26 candidate genes were grouped according to their functional categories: classical virulence factors, cell surface components, stress response, DNA recombination and repair, transport, metabolism, regulation and hypothetical genes. This STM study resulted in the identification of known virulence genes such as the *cdtA*, *hgbA* and *lspA1* which validated the STM approach. The identification of a heat shock protein and genes involved in DNA repair, emphasize the importance of stress response genes *in vivo*.

Many novel genes were discovered that had not been previously identified as having a role in pathogenesis. Genes involved in metabolism were identified which highlight the fitness contribution these factors play within a host. Two of these genes (*frdA* and *sucA*) may facilitate the ability of *H. ducreyi* to grow anaerobically. As in other STM studies, genes with hypothetical functions were identified. One gene, *hicB*, is suggested to encode an RNA toxin-antitoxin cassette. This study provides proof-of-principle for the STM approach for the recovery of genetic determinants in *H. ducreyi* responsible for the pathogenesis of chancroid.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. B. Craig Lee, who has been an excellent mentor. You have provided continuous guidance, encouragement and support. You were forever the optimist when needed and believed in me from the start. You have taught me invaluable lessons that I will never forget and will take with me in the future. I would like to thank Dr. D. Bill Cameron for his generosity and constant reassurance. You have passed on the important concept that all problems are caused by infectious diseases. I would like to thank my past committee members Dr. JR Dillon and Dr. J Webb for their advice in the beginning. I would like to thank my present committee members Dr. T Mah and Dr. A Stinzi for their help to see it through to the end. A warm thank-you goes to Dr. M. Desjardins for teaching me the bunny model and passing on his *H. ducreyi* knowledge.

Many thanks go out to my friends, colleagues and staff at Roger Guindon who have helped in more ways than they know. Thank-you to Charlene, Fiona, Chris and Nicole, without them, the journey would have been unbearable. The time spent there was filled with many conversations, laughs and tears. I am indebted to your friendship and compassion. A heartfelt thank-you goes out to the friends I have made in the Lee/Cameron/Garber lab who were there everyday and helped whenever they could. Keely, Cathy, Julie, and Jeff, you made the lab a joy and an interesting place to be everyday.

Finally, a special thank-you goes to my family for being the foundation of my hopes and dreams. I am grateful for your endless strength and patience. You are my inspiration, my comfort and my happiness.

TABLE OF CONTENTS

List of Figures

List of Tables

List of Abbreviations

ACVS	Animal Care and Veterinary Services
Ap	Ampicillin
ATCC	American Type Culture Collection
bp	base pair
CA	Chocolate Agar
CDC	Centers for Disease Control
CDT	Cytolethal Distending Toxin
CFU	Colony Forming Units
CFU/mL	Colony Forming Units per milliliter
CI	Competitive Index
Cm	Chloramphenicol
ddH_2O	Double Distilled Autoclaved Water
DIG	Digoxigenin
DNA	Deoxyribonucleic Acid
DNase	Deoxyribonucleic
DNS	Data Not Shown
DsrA	Ducreyi Serum Resistance protein A
EDD	Estimated Delivered Dose
EIA	Enzyme Immunoassay
EtOH	Ethanol
F	Farad
g	Gravitational Force

GC	Gonococcus Culture Media
GI	Gene Identification
h	Hour
H&E	Hematoxylin and Eosin
HgbA	Hemoglobin Binding receptor A
HIV	Human Immunodeficiency Virus
HP	Human Passaged
int	Integrase
IPTG	Isopropyl-beta-D-thiogalactopyranoside
Kan	Kanamycin
kb	Kilobase
kJ/sq.cm	Kilojoule per square centimeter
kg	Kilogram
LB	Luria-Bertaini
LD_{50}	Lethal Dose that kills 50% of the animals tested
LOS	Lipooligosaccharide
LPS	Lipopolysaccharide
M	Molar
m^2	Meter Squared
MB	Megabase
min	Minute
mL	Milliliter
mM	Milimolar

mRNA	Messenger Ribonucleic Acid
msec	Millisecond
nm	Nanometer
OD	Optical Density
OHRI-OGIC	Ottawa Hospital Research Institute Ontario Genomic Innovation Center
OMP	Outermembrane Protein
ORFs	Open Reading Frames
PCR	Polymerase Chain Reaction
RNA	Ribonucleic Acid
RNase	Ribonuclease
rpm	Revolutions Per Minute
rRNA	Ribosomal Ribonucleic Acid
RT	Room Temperature
SCOTS	Selective Capture Of Transcribed Sequences
sec	Second
ShlA	*Serratia* Hemolysin protein A
SPI-2	*Salmonella* Pathogenicity Island 2
ST	Signature Tag
STM	Signature-tagged Mutagenesis
STM-TDRM	Signature-tagged Mutagenesis in the Temperature-dependent Rabbit Model
TAS	Toxin-antitoxin System
TdhA	TonB-dependent Heme receptor
TDRM	Temperature-dependent Rabbit Model

T_m	Melting Temperature
U	Unit
µg	Microgram
µL	Microliter
UNAIDS	United Nations Joint Program on HIV/AIDS
UV	Ultraviolet
V	Volt
Van	Vancomycin
v/v	Volume per Volume
WHO	World Health Organization
w/v	Weight per Volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D galactopyranoside
Xis	excision

CHAPTER 1: INTRODUCTION

1.1 *Haemophilus ducreyi*:

1.1.1 *Haemophilus ducreyi*:

Haemophilus ducreyi is a Gram-negative bacterium that causes the sexually transmitted genital ulcer disease, chancroid and was first described in 1189 by Auguste Ducrey (6). Through rRNA analysis, *H. ducreyi* has now been classified in the genus of *Pasteurellaceae* that colonizes genital mucosal surfaces and shares many of the same surface antigens. It was originally grouped with *Haemophilus* species because of biochemical properties and growth requirements but rRNA analysis demonstrated that it was not that closely related to members of *Haemophilus* species but had more similarities in the rRNA cistrons of the *Actinobacillus* cluster (36). *H. ducreyi* is an unencapsulated fastidious coccobacilli that has an absolute requirement for heme to grow (6). Free hemin is not required because *H. ducreyi* can utilize hemoglobin, myoglobin and certain other heme proteins, such as catalase, as a source of heme (76). In microscopy, a Gram's stain of *H. ducreyi* clinical swabs reveals a characteristic "school of fish" pattern as observed as pairs of bacteria in short or long chains when grown on solid media and "railroad tracks" of parallel chains or large aggregates when grown in liquid media. It has a length that ranges from 1 to 2 microns and a width usually around 0.5 microns with rounded ends. The genome of *H. ducreyi* 35000HP has been sequenced and is composed of a single 1.7 MB chromosome with 1 693 putative open reading frames (ORFs). *H. ducreyi* is a strict human pathogen with no known animal or environmental reservoir (92).

1.1.2 Chancroid Infection and Clinical Features:

H. ducreyi is believed to cause infection in a host through small abrasions in the epithelium that occur during sexual intercourse (92). The clinical progression of chancroid, or soft chancre, to form an ulcer in four stages takes on average seven days but can vary. The first stage is redness or erythema at the entry site that goes on to form a papule (second stage). The papule evolves into a pustule at the third stage, which leads to the final stage of ulceration (6). The tender ulcers that form are usually well-defined with ragged edges and are not indurated. The base of the ulcer is covered by necrotic, purulent exudate that bleeds when scraped. Patients do not usually seek medical attention until they have painful ulcers, which can be complicated by inguinal lymphadenopathy and bubo formation (6). In males, lesions are usually localized to preputial mucosal surfaces or stratified squamous epithelium on the prepuce, frenulum and the coronal sulcus, although the lesions can occur anywhere. In women, the lesions are usually found at the vaginal entrance but internal lesions are also formed. Untreated ulcers can persist for 1-3 months before resolution. *H. ducreyi* does not disseminate systemically but extragenital and adjacent lesions can occur and are thought to be due to autoinoculation (146). Women are thought to be asymptomatic carriers of *H. ducreyi* because more males seem to present themselves with the infection but disease-free carriage is uncommon or rare (92).

1.1.3 Epidemiology:

The World Health Organization (WHO) and The United Nations Joint Program on HIV/AIDS (UNAIDS) estimates that the annual incidence of chancroid is around six to seven million cases worldwide, which may be a conservative estimate because many cases

go unreported or undiagnosed due to scarce resources and limited diagnostic tools (131). Particularly in developing countries, a clinical diagnosis may not be feasible due to the limited capacity of health care clinics to perform laboratory tests. Also the diagnosis of chancroid ulcers can be misinterpreted because the lesions may resemble those of syphilis or herpes. Chancroid is endemic in developing countries such as those in Asia (116), Africa (25) and the Caribbean (19) and causes 23-56% of genital ulcer disease (18). In the late 1980s, chancroid used to account for up to 80% of all genital ulcers in sub-Saharan Africa (110). Currently, herpes simplex virus (HSV) is the leading cause of genital ulcer disease in Africa while the incidence of chancroid has decreased (74). This decline may be attributed to the syndromic antibacterial treatment approach to genital infections (96). The most common targets occur in people in the 18 – 45 year old age range associated with inadequate health care, drug use and/or prostitution. Although chancroid is rare in North America, there are reported outbreaks in urban areas that have high rates of drug use and prostitution (82). The number of cases reported annually in the United States has declined since 1987, from 4 986 cases to 1 399 cases in 1993 but it is likely that chancroid has also been underreported in North America (121). Some possible reasons for the underreported cases in the US were that a standard surveillance definition for clinically compatible cases was not implemented until 1990, the appropriate culture medium was not always available and clinics had limited ability to differentiate chancroid among other genital ulcer diseases such as syphilis and genital herpes (66, 82). The use of serological tests such as enzyme immunoassay (EIA) and polymerase chain reaction (PCR) in the laboratory diagnosis of chancroid are more modern methods of diagnosis but the disadvantages of these tests are the high cost and the time between collection of the specimen and the laboratory result (143). Surveillance data from

Surveillance data from the WHO and the Centers for Disease Control (CDC) suggest its reemergence as a significant sexually transmitted genital ulcer disease due to newly reported cases and the rise of human immunodeficiency virus (HIV) transmission (90).

Renewed interest in understanding the pathogenesis of *H. ducreyi* has recently arisen due to the epidemiologic association of chancroid with HIV infection (87, 96, 116). These reports indicate that *H. ducreyi* and HIV facilitate each other's transmission. There is evidence that *H. ducreyi* chancres may shed the HIV virus, increasing the risk of transmission during sexual intercourse. The presence of inflamed ulcers on the genitalia creates direct portals of entry and target cells for HIV, augmenting the probability of establishing HIV infection after exposure. HIV infected individuals have an increased risk of acquiring genital ulcers (69). *H. ducreyi* infection has been estimated to enhance HIV transmission 10- to 100-fold per sexual act (69, 79). Co-infected individuals are less responsive to the antimicrobial therapy regimen recommended for chancroid as a result of their compromised immune systems (69). Both of these mechanisms result in increased risk of HIV transmission or exposure in individuals suffering from *H. ducreyi* infection.

The recommended therapy for chancroid is to treat with oral azithromycin, ciprofloxacin, erythromycin or intramuscularly with ceftriaxone (44). Currently there are no vaccines available for chancroid and there have been reports of antibiotic resistance in *H. ducreyi* (62). Thus, there is a need to develop new and alternative antimicrobial therapies for *H. ducreyi* infection, particularly in the developing world, which may help to decrease the rate of HIV transmission through this route of exposure.

1.1.4 Pathogenesis:

The pathogenesis of *H. ducreyi* infection has not been fully elucidated. This stems, in part, from the poorly understood biology of *H. ducreyi* and the difficulties of culturing and genetically manipulating this bacterium *in vitro*. *H. ducreyi* requires rich media and a restricted ambient temperature for optimal growth. A number of potential *H. ducreyi* virulence factors have been identified. These include a soluble cytolethal distending toxin (CDT) (32), a hemolysin (102), a lipooligosaccharide (LOS) (12), pili (97), heat shock proteins (97), iron-regulated proteins or receptors (140), Cu,Zn-superoxide dismutase (Cu,Zn-SOD) (119), a hemoglobin receptor (HgbA) (42) and outer membrane proteins (OMPs) (43). Table 1 lists all of the virulence factors tested in the human challenge model of chancroid to date.

The cytolethal distending toxin of *H. ducreyi* is encoded by three genes *cdtA*, *cdtB* and *cdtC* and was found in a majority of the strains tested (78, 113). Much of the understanding of the secreted CDT mechanism is based on *in vitro* experiments on cultured human and mammalian cells where it causes cell death. The active subunit, CdtB has DNase activity that forms double-stranded breaks which initiates a signaling cascade that induces cell cycle arrest in the G2 phase (33). The exact role of CDT in the pathogenesis of *H. ducreyi* infection is not clear. A toxin-deficient mutant demonstrated that expression of CDT was not necessary in the early stages of human volunteers when it was injected and formed pustules comparable to the wild type parent (154). However, in the rabbit model of chancroid, purified CDT caused some inflammation and when injected with *H. ducreyi* caused aggravation and slow healing of the lesions (151).

Table 1. A list of *H. ducreyi* mutants tested to date in the human challenge model of chancroid.

H. ducreyi mutant (gene knockout)	Virulence phenotype in human challenge model of chancroid	Reference
Hemolysin (*hhdA*)	Mutant as virulent as wild type *H. ducreyi* 35000HP	Palmer et al 1998
Pilus (*flpA*)	Gene is required for virulence	Al-Tawfiq et al 2000
Hemoglobin A receptor (*hgbA*)	Mutant is attenuated in virulence	Al-Tawfiq et al 2000
Outermembrane protein (*ompA2*)	Gene is required for pustule formation	Throm et al 2000
Peptidoglycan associated lipoprotein (*pal*)	Gene is required for virulence	Fortney et al 2000
Ducreyi serum resistance (*dsrA*)	Gene contributes to virulence (papules were smaller than wild type)	Bong et al 2001
Cytolethal distending toxin (*cdt*)	Gene is not required for pustule formation	Young et al 2001
Superoxide dismutase C (*sodC*)	Mutant as virulent as wild type	Bong et al 2002
Tight Adhesion (*tadA*)	Gene is required for virulence	Spinola et al 2003
Ducreyi lectin A (*dltA*)	Gene is required for virulence	Leduc et al 2004
Large supernatant protein (*lspA1* and *lspA2*)	Genes are required for virulence	Janowicz et al 2004
Collagen associated (*ncaA*)	Gene is required for virulence	Fulcher et al 2006
TonB dependent heme binding protein and uncharacterized hypothetical protein (*tdXltdhA*)	Mutant as virulent as wild type	Leduc et al 2008
Enterobacterial common antigen (*wecA*)	Gene contributes to virulence (required 2.5X more mutant to cause pustule formation than wild type)	Banks et al 2008
Fibrinogen binder A (*fgbA*)	Partial attenuation (papules were smaller than wild type)	Bauer et al 2009

The role of the hemolysin in *H. ducreyi* virulence is not well defined. Expression of the cell-associated hemolysins are encoded by *hhdB* and *hhdA* (102). These genes are similar to the hemolysins, ShlA and ShlB of *Serratia marcescens* which lyse targets cells. In *S. marcescens*, ShlB is an outer membrane protein that activates ShlA which interacts with target cell membranes and forms 3nm pores (120). *H. ducreyi* hemolysin has shown cytotoxic effects on human foreskin fibroblasts, erythrocytes, foreskin epithelial cells and immune system cells such as macrophages, T cells and B cells (152). However, in the human challenge model of *H. ducreyi* infection, an isogenic hemolysin-deficient mutant was not required for pustule formation because pustules formed at rates similar to the wild type parent strain *H. ducreyi* 35000 (104). The role of the hemolysin in the later stages of infection has not been determined.

The lipooligosaccharides (LOS) of *H. ducreyi* are different from other Gram negative enteric lipopolysaccharides (LPS) because they lack the repeating O-antigens (85). Injection of *H. ducreyi* LOS produces skin abscesses in rabbits which were larger in size to those produced from *E. coli* LPS. This suggests that LOS may play a role in its pathogenesis.

The pili found on *H. ducreyi* appear as fine tangled surface appendages that are morphologically different from the pili found on *Neisseria gonorrhoeae* (129). Pili were found in all 12 strains that were tested. A pilus preparation that was purified from *H. ducreyi* 35000 was used in a booster immunization procedure in the temperature dependent rabbit model (TDRM) of chancroid infection. Immunization with the pilus preparation significantly reduced the severity of the disease and the duration of infection compared to the controls (38). Immunization with the pilus preparation led to partial protection and

accelerated healing of the rabbits (38). However, injection of the pilin preparation in human challenge subjects demonstrated that it was not required for virulence. The pilin receptor on host cells has not been identified leading to an unknown mode of action.

The heat shock operon of *H. ducreyi, groE,* has two genes *groES* and *groEL* that show homology with the *groE* genes found in other bacteria. Northern blots of *H. ducreyi* showed high levels of *groE* mRNA when the temperature was changed from the optimal 33°C to 37°C. This suggests that *groE* had a protective effect on cells during the temperature shift and that these genes are required for *H. ducreyi* to survive and cause disease at its optimal temperature (106).

The Cu, Zn-superoxide dismutase of *H. ducreyi* is a heme-binding protein found in the periplasm (101, 119). This protein is thought to protect the bacterial cell by scavenging superoxide radicals derived from periplasmic or environmental sources such as host immune cells. The heme-binding component of this protein may play a role in heme transport, metabolism and detoxification of this organism (101).

The heme requirement of *H. ducreyi* suggests that the proteins involved in this process would be crucial to the survival of the pathogen. Two important heme proteins that have been discovered are HgbA and TdhA. HgbA and TdhA are outer-membrane proteins that share homology with other TonB-dependent receptor proteins, which are involved in heme or iron acquisition and transport (42, 140). An *hgbA* mutant is unable to bind or utilize hemoglobin as a source of heme (4). The *hgbA* mutant was shown to be attenuated in a host because the lesions produced in the rabbit model were much smaller compared to the wild type parent *H. ducreyi* 35000 (135). In the human experimental model of *H. ducreyi* infection, the HgbA-deficient mutant was attenuated and unable to form pustules on the

upper arms of human volunteers but the TdhA-deficient mutant was just as virulent as the wild type parent (4, 75). The actual mechanism of these proteins and their role in *H. ducreyi* infection still remains to be elucidated.

A number of outermembrane proteins in *H. ducreyi* have been discovered. As surface exposed molecules these proteins could serve as potential targets for bactericidal agents. *H. ducreyi* expresses two porin proteins, OmpP2A and OmpP2b that have been shown to be determinants in serum resistance. A double mutant was injected into the upper arms of human volunteers and demonstrated that it was not necessary for pustule formation because pustules formed at the same rate as the wild type parent strain (64). A 30 kDa outer membrane protein DsrA has been discovered in most virulent strains of *H. ducreyi* to be required for serum resistance (43). An isogenic *dsrA* mutant was tested in the human model of *H. ducreyi* infection and was shown to have decreased virulence. The mutant developed smaller papules than its wild type parent and was impaired in its ability to progress to the pustular stage (20). The bactericidal activity of human serum is an important component of innate host defenses, and many bacterial pathogens are serum resistant. Thus, these proteins suggest possible mechanisms to target for treatment.

Many of these putative virulence determinants of *H. ducreyi* have been discovered using *in vitro* techniques that may not reflect complex host-parasite interactions. Their roles *in vivo* are also not well defined. Consequently, genes that are regulated or activated by the host may not be discovered through *in vitro* methods and their importance may be overlooked.

The ultimate objective of a pathogen is to survive and replicate by first gaining access to and then persisting within host sites. These interactions within a host require the

pathogen to not only replicate but they may be associated with pathological lesions and symptoms of disease. The control of gene expression in bacteria has been proposed to be an adaptation in response to environmental cues. Temperature (73, 125), pH (17, 99), oxygen tension (3), nutrient limitation (7) and cation concentration (49) induce a variety of virulence factors *in vitro.* In this context, the definition of virulence factors has been expanded to include not only classical virulence determinants such as secreted toxins but also factors that are necessary for the fitness and survival of the pathogen within a host and for its transmission to new hosts. The expansion of the definition emphasizes the importance of ecological factors, differences in host clearance mechanisms and evolutionary factors that are involved in the transmission efficiency of the microbe to new susceptible hosts (136). For bacterial pathogens, the host milieu is thought to enforce a bias on bacterial gene regulation by selective expression of gene products that are important for the survival and infection of the pathogen within the host. Virulence factors can also be coordinately controlled by global regulons in response to changes of different host conditions even if individual genes do not make direct a contribution. Thus, the most appropriate method to identify and examine bacterial virulence determinants would be through *in vivo* analysis.

1.1.5 Animal Models:

There are a number of animal models of infection for chancroid such as a swine ear model (57), a macaque model (144), a human forearm challenge model (130) and a temperature-dependent rabbit model (TDRM) of infection (112).

In the human infection model of chancroid, *H. ducreyi* is injected into the epidermis and dermis of the upper are by the tines of an allergy testing device (130). The estimated

delivered dose (EDD) that is required to initiate papule formation is between 1 to 100 CFU. After 24 hours of inoculation, papules have formed and development into pustules takes 2 to 5 days. In some cases, the pustules resolve and do not develop any further. The end point of this challenge model is determined by painful pustules or after 14 days of infection. For safety precautions, the pustules in the human challenge model only reach stage 2 of development and are not allowed to ulcerate (130). Antibodies are not detected within the 2 weeks of the model (103). However, in naturally occurring *H. ducreyi* infection, antibodies develop after 3 weeks of ulceration (26). A limitation of this model is that it can only be used to study the first two weeks of an infection and cannot be used to study ulcers.

In the swine model, *H. ducreyi* is injected into the ears of pigs using an allergy testing device similar to the one used in the human challenge model. The EDD of bacteria for this model is 10^4 CFU (57). Within 2 days of inoculation, papules start to form and pustules develop by 7 days. Ulceration occurs by 14 days of infection. Antibodies develop to *H. ducreyi* antigens within 1 to 2 weeks of infection. The major draw back of this model is the cost of the animal and that the large size of the animal requires adequate facilities to house and monitor it.

The *H. ducreyi* macaque infection model develops ulcers within 1 to 2 days after intraepithelial injection of 10^7 CFU bacteria into the foreskin of males (144). Inguinal lymphadenopathy can develop and the lesions that form in this model most closely resemble chancroidal ulcers in human patients. Antibodies to *H. ducreyi* antigens were also detected within 2 weeks of inoculation. However, female primates injected in the genitalia with *H. ducreyi* do not develop lesions (144). Therefore this model can only be used to assess *H. ducreyi* infection in males.

Each model has inherent limitations from the development of persistent lesions such as in the macaque model, to the duration of the infection such as in the human challenge model that does not develop to the ulcerative stage and also the economical feasibility of the swine and macaque models. The best described model of *H. ducreyi* infection is the TDRM in which rabbits are housed at 15°C to 17°C to mimic the lower temperature of the human genital areas where the ulcers develop. The TDRM has been developed as both a quantitative and qualitative model (38). It has been used to evaluate the virulence of *H. ducreyi* isolates and to demonstrate that immunization can induce protective immunity against experimental chancroid (37, 53, 78). The use of *in vivo* models such as the TDRM may be the key in understanding the mechanisms of how *H. ducreyi* invades its host and establishes an infection (112). This model reflects the host's role in bacterial gene regulation and induction of specific genes required for survival and virulence of *H. ducreyi* within a host. The identification of new virulence genes may enable the development of novel therapies and strategies to treat or prevent chancroid.

1.2 Signature-tagged Mutagenesis (STM):

Recent advances in genomic technologies have created new *in vivo* assays to identify potential survival and virulence genes that are activated through pathogen-host interactions (as reviewed in (9, 28, 83, 117)). Signature-tagged mutagenesis (STM) is an *in vivo* mutation-based genetic screen that allows for the large-scale identification of microbial virulence genes activated upon infection with a host (147). STM uses a pool of tagged insertion mutants to inoculate an animal. The animal is used to isolate mutants that are

unable to survive the specified environmental conditions in the host and was first used to identify novel virulence genes in *Salmonella typhimurium* (55).

STM has proven to be a powerful and versatile global genetic technique to identify virulence determinants. The first STM screen performed more than 10 years ago in *S. typhimurium* discovered a novel pathogenicity island (SPI-2) whose genes were involved in host cell invasion (55). STM has since been used to identify novel virulence genes in a number of other organisms and has been tested in over 40 STM screens, many of which do not have well defined genetic systems to manipulate (83). Since then, other STM screens have been tested in many other organisms including *Vibrio cholerae* (27), *Mycobacterium tuberculosis* (23), *Yersinia enterocolitica* (35), *Escherichia coli* K1 (11), *Aspergillus fumigatus* (22), *Brucella suis* (47), *Listeria monocytogenes* (10), *Toxoplasma gondii* (72), *Cryptococcus neoformans* (95), *Burkholderia pseudomallei* (8) and *Xenorhabdus nematophilia* (56).

From these studies a wide variety of putative virulence genes have been discovered including genes involved in invasion of host cells, DNA replication, regulation, cellular metabolism, cellular repair and transport as well as previously known virulence factors (108). The identification of previously known virulence determinants validates the STM approach. However, only a small percentage of the genes identified in STM studies are considered classical or genuine virulence genes such as adhesions, invasions, toxins and effectors. In many STM screens, insertions in genes that encode surface molecules were identified which showcase the importance of these proteins in pathogenesis. Surface molecules are important for many processes such as cell tropism, crossing barriers, stress adaptation, colonization and dissemination. Genes involved in the biosynthesis or

modification of peptidoglycan cell wall components, in addition to adhesions and transporters have been identified. Membrane transporters are important elements in metabolite trafficking. The identification of genes involved in biosynthetic pathways and metabolism demonstrates that these genes are necessary to fulfill essential nutritional requirements that need to be met in different environments. STM studies have identified regulatory genes that are essential for expressing genes in changing environments highlighting their importance in developing and maintaining the infection. STM has also uncovered many potential virulence factors that show no homology to any genes within the available databases and have been classified as genes with unknown function (108). Many of the genes identified have not been experimentally shown to be involved in pathogenesis. Thus the definition of a virulence factor has been expanded to not only include classical genes such as toxins but to include all genes that are important for the survival and persistence of a pathogen in a host.

The STM approach relies on the animal model to negatively select attenuated bacterial strains from a library of transposon-derived mutants. The important feature of this technique is that each transposon contains a unique signature tag that allows the fate of each mutant to be monitored (147). These genes can then be assessed and characterized for their survival and virulence role. The technique is outlined briefly below (Figure 1). A mixture of tagged mutants is used to establish an infection in an animal model. The mutants that are initially present are later compared to the mutants that remain at the death of the host. Those lost during the infection process are presumed to be attenuated as a result of the insertion of the transposon into a gene that is required for survival in the host. Because the tagged transposons carry an antibiotic resistance marker, the region flanking the transposon can be

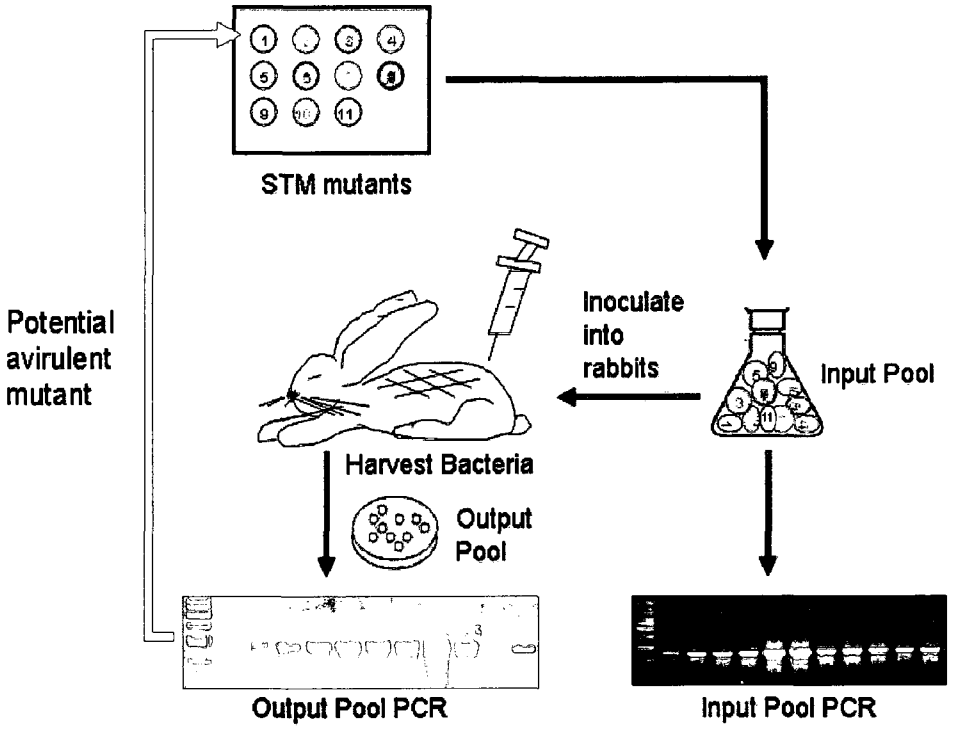
subcloned from a restriction digest using antibiotic selection. The subcloned gene carrying the insertion can be identified by using the tag as a primer to sequence the DNA flanking the transposon and comparing the sequence to known bacterial DNA databases.

1.2.1 PCR-based STM:

In the original STM approach to identify virulence genes in *Salmonella typhimurium*, the tags consisted of short DNA segments containing a 40 bp variable central tag that was flanked by invariant 20 bp regions (55). The suitability of individual tags for detection within a pool of differently tagged mutants was tested by pooling single mutants and testing for hybridizations. The tags were PCR amplified with universal tag primers, radioactively labeled and hybridized on membranes spotted with DNA from the corresponding mutants. Only mutants with clear tag hybridization signals were included in the subsequent selection process. After the input pools of mutants were subjected to negative selection in an animal model, the corresponding output pools recovered were grown on media and their tags amplified and labeled for detection. A weak or absent hybridization signal from the output pool compared to the input pool identified attenuated mutants.

An alternative to hybridization as the selection method was created in which the tags of the mutants were analyzed directly by PCR amplification to indicate the presence of absence of the tags in the *Pseudomonas aeruginosa* rat lung model of infection (77). This method relies on using primers that are specific for each tag. The advantage of this STM technique is that the hybridization step is circumvented. However this approach is less quantitative and a large number of PCR products must be analyzed by gel electrophoresis. PCR-STM has been used to identify new virulence determinants in

Figure 1. A PCR-based signature-tagged mutagenesis schematic. The red box indicates a mutant that was not detected by PCR in the output pool suggesting a potential avirulent mutant.


STM mutants
Inoculate into rabbits
Input Pool
Potential avirulent mutant
Harvest Bacteria
Output Pool
Output Pool PCR
Input Pool PCR

S. gallinarum (122), *Francisella tularensis* (137), *Edwardsiella ictaluri* (141), *Actinobacillus suis* (98) *and Helicobacter pylori* (70). Two of the studies (*A. suis* and *H. pylori*) expanded the original number of PCR STM tags from 12 to 24 to generate more complexity and mutants in each pool to screen in their respective animal models. The STM screen of *E. ictaluri* in catfish identified virulence genes that were related to the SPI-2 Type III secretion system and in the *H. pylori* infection of a gerbil, a secreted collagenase was identified as a potential virulence factor (70, 141). These STM screens identified putative genes to characterize and provided new avenues of study towards the understanding of virulence mechanisms and host restrictions of these pathogens (Figure 1).

1.3 Research Hypothesis:

We hypothesize that the expression of genes that are crucial for the pathogenesis of infection due to *H. ducreyi* are regulated by environmental stimuli induced upon infection of a host. In the absence of such genes, *H. ducreyi* would be less virulent. Strategies designed to identify and interfere with the function of these genes may provide novel therapies to combat chancroid, which can also be used as preventative therapy against HIV infection.

1.3.1 Research Objectives:

Our objective is to apply the STM technique to investigate the *H. ducreyi* genome for virulence genes activated upon infection of a host. The basic steps of STM are the following: (1) construction of signature-tagged transposons, (ii) generation of tagged transposon libraries, (iii) inoculation into an animal model, (iv) recovery of attenuated mutants, (v) comparison of *in vitro* and *in vivo* pools, and (vi) phenotypic characterization of mutants (147). In order to do this, the research objectives were:

(1) Construction of signature-tagged transposons

(2) Generation of signature-tagged transposon mutant libraries

(3) Applying STM to *H. ducreyi* in the TDRM and screening of the mutants

(4) Sequence and reconfirm the attenuation of candidate ST *H. ducreyi* mutants

CHAPTER 2: MATERIALS AND METHODOLOGY

2.1 Bacterial Strains and Growth Conditions:

Haemophilus ducreyi 35000 HP 700724 was obtained from American Type Culture Collection (ATCC; Manassas, VA). The *H. ducreyi* 35000 strain was originally isolated in the 1975 Winnipeg chancroid outbreak and was used for signature-tagged mutagenesis experiments. *H. ducreyi* was grown at 35°C in an atmosphere of 5% CO_2 and saturated humidity.

Escherichia coli TOP10 cells were purchased from Invitrogen (Carlsbad, CA) and were used as the host cell for all intermediate DNA manipulations and cloning (Table 2). *E. coli* strains were grown at 37°C in ambient air.

Oligonucleotide primers and plasmids used in this study are listed with their relevant properties in Table 2 and Table 3 respectively.

2.2 Media and Culture Conditions:

Haemophilus ducreyi 35000 HP 700724 was stored in freezing media [25% (v/v) glycerol, 3% (w/v) Trypticase Soy Broth (Gibco/BRL, Rockville, MD)]. Stocks were maintained at -70°C. When required, aliquots were grown on chocolate agar [CA; GC agar base (Difco, BD; Franklin Lakes, NJ) with 1% (w/v) bovine hemoglobin (BBL/Becton Dickinson, Cockysville, MD), supplemented with 5% (v/v) fetal bovine serum (FBS; Gibco/BRL, Rockville, MD) and 1% (v/v) IsoVitaleX (BBL/Becton Dickinson, Cockysville, MD)] at 35°C in 5% CO_2 for 24 h at 175 rpm. GC broth [1.5% (w/v) proteose peptone, 23 mM K_2HPO_4, 7.3 mM KH_2PO_4 (both from Fisher Scientific, Fair Lawn, NJ), and 86mM NaCl (BDH, Toronto, ON)], supplemented with 5% (v/v) FBS and 1% (v/v) IsoVitaleX was inoculated from overnight agar plates and the liquid

Table 2. Oligonucleotide primers used in this study.

Primer Name	Sequence (5'—> 3')	Function
Tnl545K-F	AGCGAACCATGGTACCTTGAGGTGAT	Amplify Tnl545-Δ3
Tnl545K-R	GTTTTCCCAGGGTACCTCACGACGTT	Amplify Tnl545-Δ3
F-Kan	AATTGCCACACATCACTCCA	*Aph3* cassette Tnl545-Δ3
R-Kan	CCAATTCACTGTTCCTTGCAT	*Aph3* cassette Tnl545-Δ3
Xis-R	CGCTCCTGTTGCTTCTCTTT	*Xis & Int* genes of Tnl545-Δ3
Int-L	AATTGCCACACATCACTCCA	*Xis & Int* genes of Tnl545-Δ3
STMlinkR2	GAACCTGCTAGCAGATCTTTGCCGATTCATAACAGCCA	STM linker
STM1linkF3	GTTAATAGATCTGTACCGCGCTTAAACGTTCAGGAAAGG AGCAAATGCCATGT	STM1 linker
STM2linkF	GTTAATAGATCTGTACCGCGCTTAAATAGCCTGGAAAGG AGCAAATGCCATGT	STM2 linker
STM3linkF	GTTAATAGATCTGTACCGCGCTTAAAAGTCTCGGAAAGG AGCAAATGCCATGT	STM3 linker
STM4linkF	GTTAATAGATCTGTACCGCGCTTAATAACGTGGGAAAGG AGCAAATGCCATGT	STM4 linker
STM5linkF	GTTAATAGATCTGTACCGCGCTTAAACTGGTAGGCCCGG AGCAAATGCCATGT	STM5 linker
STM6linkF	GTTAATAGATCTGTACCGCGCTTAAGCATGTTGGAAAGG AGCAAATGCCATGT	STM6 linker
STM7linkF	GTTAATAGATCTGTACCGCGCTTAATGTAACCGGAAAGG AGCAAATGCCATGT	STM7 linker
STM8linkF	GTTAATAGATCTGTACCGCGCTTAAAATCTCGGGAAAGG AGCAAATGCCATGT	STM8 linker
STM9linkF	GTTAATAGATCTGTACCGCGCTTAATAGGCAAGGAAAGG AGCAAATGCCATGT	STM9 linker
STM10linkF	GTTAATAGATCTGTACCGCGCTTAACAATCGTGGAAAGG AGCAAATGCCATGT	STM10 linker
STM11linkF	GTTAATAGATCTGTACCGCGCTTAATCAAGACGGAAAGG AGCAAATGCCATGT	STM11 linker
STM12linkF	GTTAATAGATCTGTACCGCGCTTAACTAGTAGGGAAAGG AGCAAATGCCATGT	STM12 linker
STM1	GTACCGCGCTTAAACGTTCAG	STM1 tag
STM2	GTACCGCGCTTAAATAGCCTG	STM2 tag
STM3	GTACCGCGCTTAAAAGTCTCG	STM3 tag
STM4	GTACCGCGCTTAATAACGTGG	STM4 tag
STM5	GTACCGCGCTTAAACTGGTAG	STM5 tag
STM6	GTACCGCGCTTAAGCATGTTG	STM6 tag
STM7	GTACCGCGCTTAATGTAACCG	STM7 tag
STM8	GTACCGCGCTTAAAATCTCGG	STM8 tag
STM9	GTACCGCGCTTAATAGGCAAG	STM9 tag
STM10	GTACCGCGCTTAACAATCGTG	STM10 tag
STM11	GTACCGCGCTTAATCAAGACG	STM11 tag
STM12	GTACCGCGCTTAACTAGTAGG	STM12 tag
Tnl545R	GTTTTCCCACTCACGACGTT	Tnl545-Δ3
HduHgbAF	CCGAAAGCAATATGCAAACA	Amplify *HgbA*
HduHgbAR	TCATCCTAAATGACCTAGTTAGCC	Amplify *HgbA*
STMkanF	ATTCAACGGGAAACGTCTTG	Nested inverse PCR
STM1R	CTGACCGTTTAAGCGCGGTAC	Nested inverse PCR
pMS1kanR	CTTTTCCGACGCATTTATCG	Nested inverse PCR
pMS1xisF	AGAATTGGCTTCATAACCATTG	Nested inverse PCR

Table 3. Bacterial plasmids and constructs used in this study.

Name of plasmid	Size (Kb)	Source	Function
pBluescript II KS+	3	Stratagene	Suicide vector
pMS1	9.3	Hansen et al 1995	Transposon Tn1545-Δ3
pBluTn	6.7	This study	*Tn1545-Δ3* in pBluescript
pBluTnSTM1	6.7	This study	STM1 *Tn1545-Δ3* cloning vector
pBluTnSTM2	6.7	This study	STM2 *Tn1545-Δ3* cloning vector
pBluTnSTM3	6.7	This study	STM3 *Tn1545-Δ3* cloning vector
pBluTnSTM4	6.7	This study	STM4 *Tn1545-Δ3* cloning vector
pBluTnSTM5	6.7	This study	STM5 *Tn1545-Δ3* cloning vector
pBluTnSTM6	6.7	This study	STM6 *Tn1545-Δ3* cloning vector
pBluTnSTM7	6.7	This study	STM7 *Tn1545-Δ3* cloning vector
pBluTnSTM8	6.7	This study	STM8 *Tn1545-Δ3* cloning vector
pBluTnSTM9	6.7	This study	STM9 *Tn1545-Δ3* cloning vector
pBluTnSTM10	6.7	This study	STM10 *Tn1545-Δ3* cloning vector
pBluTnSTM11	6.7	This study	STM11 *Tn1545-Δ3* cloning vector
pBluTnSTM12	6.7	This study	STM12 *Tn1545-Δ3* cloning vector
pPCR-Script Amp SK(+)	3	Stratagene	Cloning vector
pHgbA	6	This study	*HgbA* in pPCR-Script
pLS88	4	Wilson et al 1989	Native *H. ducreyi* plasmid

cultures were grown at 35°C with shaking. *H. ducreyi* transformants were grown on chocolate agar plates supplemented with 0.25% (w/v) charcoal. When necessary, chocolate agar plates were supplemented with antibiotics at the following concentrations: kanamycin (Kan) at 30 μg/ml, chloramphenicol (Cm) 1.5 μg/ml, ampicillin (Ap) 1.5 μg/ml and vancomycin (Van) 3μg/ml.

E. coli Top10 cells (Invitrogen, Carlsbad, CA) were stored in 10% (v/v) glycerol and maintained on Luria-Bertani (LB) agar or in LB broth (Difco, BD; Franklin Lakes, NJ). For induction of protein expression, X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) was added to a final concentration of 40 μg/ml. When necessary, antibiotics were added at the following concentrations: kanamycin 50μg/ml, chloramphenicol 30μg/ml, and ampicillin 100 μg/ml.

2.3 Molecular Biology Techniques:

2.3.1 Agarose Gel Electrophoresis:

Agarose gel electrophoresis was performed using electrophoresis grade agarose (Gibco/BRL, Rockville, MD) added to a final concentration of 0.8% (w/v) to 1% (w/v) in 1X TBE buffer [45 mM Tris-borate, 1 mM EDTA] or 1X TAE buffer [0.04 M Tris acetate, 0.001 M EDTA]. The electrophoresis was conducted under constant voltage ranging between 30V to 120V. The DNA bands were visualized by the addition of 0.5 μg/ml ethidium bromide (Invitrogen, Carlsbad, CA) to the melted agarose and observed under UV light using a MultiImage Light cabinet (Alpha Innotech Corporation, CA). Molecular weight markers comprised supercoiled DNA ladder, 1 kb Plus ladder, 1 kb ladder, High Mass Ladder, λ*Hind*III ladder, and 100 bp Ladder (Invitrogen Life

Technologies Carlsbad, CA). A 1 kb ladder from Fermentas (Glen Burnie, MD) was also used.

2.3.2 Polymerase Chain Reaction (PCR) Amplification:

PCR amplification using a template of either chromosomal or plasmid DNA was performed in a 25 µl reaction volume containing a final concentration of 1X PCR buffer, 0.5 U Platinum *Taq* polymerase (Invitrogen Carlsbad, CA), 1.5 mM $MgCl_2$, 0.2 mM dNTPs (Gibco/BRL, Rockville, MD), and 10 ng/µl genomic DNA or plasmid DNA as template. Thirty cycles of amplification were performed using a Genetouch gradient PCR thermocycler (Techne, Princeton, NJ). The reaction profile consisted of a 30 sec denaturation step at 94°C, a 45 sec annealing step at the appropriate temperature and a 2 min primer extension/elongation step at 72°C for 25-35 cycles. The profile for specific PCR amplification reactions was established by varying the reaction volume, the concentration of $MgCl_2$, ratio of oligonucleotide primers to template, annealing temperature and duration of primer extension. These steps were changed systematically in order to achieve the desired results. The identity of PCR products were confirmed by nucleotide sequencing.

2.3.3 Whole Cell or Colony PCR:

A single colony was selected from an agar plate and suspended in 25 µl of the PCR reaction mixture. The initial denaturing time of the PCR reaction was changed to 10 min and the subsequent steps were as described in Section 2.3.2.

2.3.4 Nested Inverse PCR:

This PCR reaction was used to help amplify the genomic DNA segment flanking the transposon in the *H. ducreyi* mutants. Ten nanograms of genomic DNA was used as a template for PCR using primers STM1R-12R and STMkanF (Table 2) with the same conditions as described above. One microliter of the PCR reaction was used as the template in a nested reaction performed under the same conditions as described above, using primers pMS1xisF and pMS1kanR (Table 2).

2.3.5 Molecular Cloning of PCR Product:

PCR amplification using chromosomal DNA as template was performed in a 25 µl reaction volume using *pfu* DNA polymerase (Invitrogen, Carlsbad, CA). DNA fragments were purified from 1% (w/v) agarose gels using the QIAquick gel extract kit (Qiagen, Alameda, CA) or the QIAgen PCR purification column (Qiagen, Alameda, CA) according to manufacturer's specifications (Section 2.3.6). The purified DNA fragments were resuspended in 25 µl of ddH_20. To add an A-overhang to the DNA fragment, a modified PCR reaction was performed. The reaction mixture contained 2.5 µg of the purified PCR product, 2.5 U *Taq* DNA polymerase (Invitrogen, Carlsbad, CA), 2.5 µl of 10X PCR buffer, 0.75 µl of 50 mM $MgCl_2$, 1 µl of 10 µM dATP, and dH_20 to a final volume of 25 µl. After incubation at 72°C for 10 min, the modified insert was ligated into the pPCR-Script vector TA cloning kit (Stratagene; Cedar Creek, TX). The insert was added at a ratio of 1:3 of vector: insert, and mixed with 0.25 µg of pPCR-Script vector in a 10 µl ligation reaction using 1U of T4 DNA ligase (Invitrogen, Carlsbad, CA). Following incubation at 4°C for 12 h, 10 µl of the ligation mixture was added to 25 µl of

competent *E. coli* Top10 cells. The sample was then incubated for 30 sec at 42°C, followed by the addition of 1 ml of SOC media [2% (w/v) Bacto Tryptone, 0.5% (w/v) Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10mM $MgSO_4$, 20 mM glucose]. After a further incubation at 37°C for 1 h with shaking at 225 rpm, cells were transferred onto LB (Ap and/or Kan) agar plates and incubated overnight at 37°C.

2.3.6 Purification of DNA Fragment from Agarose Gels:

DNA fragments obtained by PCR amplification were separated on a 1% (w/v) agarose gel. The DNA fragments were recovered from the gel either by the QIAquick Gel Extraction Kit (Qiagen, Alameda, CA), the QIAgen PCR Purification kit (Qiagen, Alameda, CA) or the Ultrafree-DA Filter Unit (Millipore Corporation, Bedford, MA, USA) according to the manufacturers' protocol. The DNA samples were resuspended in 25 to 50 µl of ddH_20. Gel extracted DNA fragments were electrophoresed on an agarose gel to confirm purity and sizes.

2.3.7 Cracking Assay:

Transformants were selected from agar plates, and placed into 1.5 ml polypropylene tubes containing 25 µl ddH_20. 25 µl of Cracking solution [1 M NaOH: 2X Cracking buffer (0.5 M EDTA, pH 8.0, 0.1% (w/v) SDS, 1% (v/v) glycerol, 0.005% (w/v) bromophenol blue)] was added to each tube. Tubes were incubated at room temperature (RT) for 5 min. A 10 µl aliquot from each sample was analyzed by electrophoreses on a 0.8% (w/v) agarose gel. Transformants containing the appropriate

sized plasmid were inoculated into antibiotic containing LB broth and grown for 12 h at 37°C.

2.3.8 Preparation of Plasmid DNA:

A 5 ml aliquot from an overnight broth culture of *E. coli* or *H. ducreyi* was pelleted by centrifugation at 4000 x g in a Juan Benchtop Centrifuge model CR3i (Saint-Herblain, France) in a 15 ml Falcon tube (Fisher Scientific, Fair Lawn, NJ) for 10 min at 4°C. Plasmid DNA was extracted using a QIAgen miniprep kit (Qiagen, Alameda, CA) as per the manufacturer's instructions. The DNA was resuspended in 50 µl of either10 mM Tris-HCl pH 8.5 or ddH_20. Plasmid DNA was quantified using a Genequant II DNA calculator (Amersham Pharmacia Biotech, Piscataway, NJ) or by comparison of band intensity to a commercially available High Mass ladder (Invitrogen, Carlsbad, CA) on an agarose gel.

2.3.9 Preparation of Genomic DNA for Automated DNA Sequencing:

A 10-15 ml aliquot from an overnight broth culture of *E. coli* or *H. ducreyi* was pelleted by centrifugation at 4000 x g in a Juan Benchtop Centrifuge in a 50 ml falcon tube (Fisher Scientific, Fair Lawn, NJ) for 15 min at 4°C. Genomic DNA was extracted using a Qiagen Genomic extraction kit (Qiagen, Alameda, CA) as per the manufacturer's specifications. The DNA was resuspended in 20 µl ddH_20. Genomic DNA was quantified using a Genequant II DNA calculator (Amersham Pharmacia Biotech, Piscataway, NJ) or by comparison of band intensity to a commercially available High Mass ladder (Invitrogen, Carlsbad, CA) on an agarose gel.

2.3.10 DNA Sequencing:

Nucleotide sequencing was performed by the University of Ottawa Core Sequencing and at the Ottawa Hospital Research Institute Ontario Genomic Innovation Centre (OHRI-OGIC). Cloned flanking DNA was sequenced using Big Dye Terminators v 3.1 Chemistry according to the manufacturer's directions and sequence data was generated on an Applied Biosystems 3730 DNA analyzer (ABI Biosystems, Columbia, MD). Sequence alignments and analyses were performed using MacVector (Oxford Molecular, Palo Alto, CA). Similarity searches of GenBank were performed by BlastN (DNA sequences) and BlastX (protein sequences). Searches of the *H. ducreyi* genome Project Database were performed by BlastN.

2.3.11 Restriction Enzyme Digests:

All restriction enzymes used were purchased from New England Biolabs (NEB, Beverly, MA) and reactions were performed according to manufacturer's directions. Reaction mixtures contained 1 μg plasmid DNA, 1X NEB React buffer, 1 U restriction enzyme and ddH_20 to a final volume of 10 μl. The mixtures were incubated in a 37°C waterbath unless otherwise stated for 1-16 h. The following restriction enzymes were used: *Ava*I, *Avr*II, *Cla*I, *Kpn*I, *Nco*I, *Pst*I, *and Xmn*I.

2.3.12 Southern Blot:

2.3.12.1 Labeling of DNA Probe:

The 1 kb Digoxigenin (DIG)-labeled DNA probe was prepared as per the manufacturer's instructions using the PCR DIG labeling mix with gene specific

oligonucleotide primers F-kan and R-kan and pMS1 as template (Roche, Boehringer Mannheim, Germany).

2.3.12.2 Preparation of *H. ducreyi* Genomic DNA for Southern Blot:

Cells obtained from ten CA-Kan plates that were incubated overnight at 35°C with 5% CO_2 were suspended in a 1.5 ml microcentrifuge tube containing TE buffer pH 8.0 [10 mM Tris-HCl, 1 mM EDTA (pH 8.0)]. The cell suspension was centrifuged for 10 min at 13 000 x g at 4°C and the supernatant was decanted. The pellet was resuspended in 600 µl TE buffer pH 8.0 by vortexing. One hundred µL of lysozyme (100 mg/ml, SIGMA) was added followed by 30 min incubation at 37°C. A solution containing 33µl of 10% (w/v) SDS and 3 µl of Proteinase K (20mg/ml, Gibco/ BRL, Rockville, MD) was added and a further incubation for 30 min at 37°C was performed. The next step involved the addition of 112 µl of 5M NaCl. After mixing thoroughly by vortexing, 89 µl of 10% (w/v) CTAB (hexadecyltrimethyl ammonium bromide, SIGMA) in 0.7 M NaCl was added and the mixture was incubated for 10 min at 65°C. The protein and nucleic acid solution was then exposed to 500 µl of chloroform/isoamyl alcohol 24:1 followed by centrifugation for 5 minutes at 13 000 x g at 4°C. The aqueous phase was transferred to a clean 1.5 ml centrifuge tube and 6.7 µl RNase A (10 mg/ml) (SIGMA; Oakville, ON) was added. After 30 min incubation at 37°C, the DNA was extracted twice with 500 µl phenol/chloroform/isoamyl alcohol. The DNA was precipitated by adding an equal volume of isopropanol followed by centrifugation for 10 min at 13 000 x

g at 4°C. The DNA pellet was washed with 1 ml of 70% ethanol (EtOH) and air-dried. The DNA was resuspended in 20 µl ddH_20.

2.3.12.3 Southern Blot of Genomic DNA:

Approximately 10 µg of genomic DNA was digested with *Ava*I, *Cla*I, and *Nco*I (NEB). The digested DNA fragments were separated by a 0.6% (w/v) agarose gel in 1X TAE buffer at 30V for 16-20 h. The gel was photographed and then depurinated in 300 ml of 0.25 M HCl with gentle agitation for 10 min at RT. After rinsing with ddH_2O, the gel was incubated in 300 ml of denaturation solution [1.5 M NaCl, 0.5 M NaOH] for 45 min followed by submersion in 300 ml of neutralization solution [1.5 M NaCl, 1 M Tris pH 7.5] for 45 min at RT. The denatured DNA fragments were transferred onto a nylon membrane (PAL Biodyne; East Hills, NY) using the method by Reed et al 1985 (115). In brief, a transfer pyramid was prepared consisting sequentially from bottom to top of a glass plate for support, a wick of Whatman 3MM filter paper, the agarose gel, the nylon membrane, Whatman 3MM filter paper, paper towels or a sponge, and a glass plate. The transfer proceeded overnight in 20X SSC [3 M NaCl, 0.3 M sodium citrate, pH 7]. Following transfer, the blot was rinsed in 2X SSC and air-dried until the membrane was slightly damp. The DNA was covalently linked to the membrane by exposing the DNA side of the membrane to an ultraviolet light source for 45-60 sec at 1200 kJ/sq. cm.

2.3.12.4 Hybridization of DIG-labeled DNA Probe:

The nylon membrane was incubated with prewarmed hybridization buffer [6X SSC, 5X Denhardt's Solution, 0.5% (w/v) SDS, 0.1 mg/ml Salmon sperm DNA

(Invitrogen, Carlsbad, CA)] for a minimum of 4 h at 60°C. The gel purified DIG-labeled DNA probe was suspended in 5 ml of hybridization buffer, heated at 100°C for 5 min, cooled quickly on ice and then added to the hybridization buffer. The membrane was incubated in the hybridization solution overnight at 60°C. The blot was washed twice with 100 ml wash solution 1 [2X SSC, 0.1 % (w/v) SDS] for 5 min each at RT, followed by rinsing twice in 100 ml of wash solution 2 [0.1X SSC, 0.1% (w/v) SDS] for 15 min each at 50°C.

2.3.12.5 Chemiluminescent Detection of DIG-labeled Probe:

The blot was washed for 5 min at RT with 100 ml washing buffer [0.3% (v/v) Tween 20, 0.1 M maleic acid, 0.15 M NaCl, pH 7.5] followed by an incubation at RT in 100 ml of blocking buffer [1% DIG blocking solution,0.1 M maleic acid, 0.15 M NaCl, pH 7.5] for 30 min. The DIG-AP-conjugate was diluted 1:5000 in 20 ml of blocking buffer and reacted with the membrane for 30 min at RT. The membrane was washed twice with 100 ml of washing buffer for 15 min each at RT, followed by equilibration in 20 ml detection buffer [0.1 M Tris, 0.1 M NaCl and 0.5 M $MgCl_2$, pH 9.5]. The blot was developed in a 1:100 dilution of AMPPD [3-(2'spiroadamantane)-4-methoxy-4-(3'phosphoryloxyl)-phenyl-1, 2-dioxetane] for 5 min at RT. After the excess AMPPD solution was allowed to drain from the membrane, the blot was sealed in a plastic bag, and incubated for another 15 min at 37°C. The membrane was placed in an X-ray film cassette to be exposed to BioMax Light Chemiluminescence film (Sigma-Aldrich; Oakville, ON) from 1 min to several hours, in order to optimize the band to background intensity ratio. The film was processed in an x-ray film developer.

2.4 Transformation:

2.4.1 Transformation of *E. coli*:

A tube containing 25 µl of competent *E. coli* Top10 cells was removed from the -70°C freezer and placed on ice to gently thaw. An appropriate amount (10 - 50 ng) of the plasmid DNA was added to the cell suspension followed by gentle mixing by inversion. The reaction mixture was incubated on ice for 30 min followed by a 30 sec exposure at 100°C at 42°C. After the addition of 500 µl of SOC media, the cell mixture was incubated at 37°C for 1 h with shaking at 225 rpm in a console orbital shaker (Forma Scientific, Marietta Ohio). A 100 µl volume of the transformed cells were plated on antibiotic containing LB agar and the plates were incubated overnight at 37°C. Agar plates with 40 µl of 2% X-gal and/or 40 µl 10 mM IPTG (isopropyl-beta-D-thiogalactopyranoside) in addition to the antibiotic(s) were used when cloning into the multiple cloning site of pBluescript or pPCR-Script to facilitate screening of transformants.

2.4.2: Transformation of *H. ducreyi*:

H. ducreyi HP 35000 was transformed by electroporation according the protocol established by Totten et al 1995 (145) with minor modifications. Cells from five CA plates that were grown overnight at 35°C in an atmosphere of 5% CO2 were suspended in a 1.5 ml centrifuge tube containing 10% (v/v) sterile cold glycerol in ddH_2O. The cell suspension was centrifuged at 13 000 x g for 2 min, the supernatant decanted and the pellet resuspended in 10% (v/v) glycerol in ddH_2O. This step was repeated twice, and the pellet was resuspended in an equal volume of 10% (v/v) glycerol in ddH_2O.

A 40 µl aliquot of bacteria was mixed with 1 µg of plasmid DNA in a 0.1 cm electrode gap cuvette (BioRad). Samples were electroporated using the BioRad Micro Pulser at a setting of 2.5 V, 200 Ω and 3.0 F for a minimum of 5 msec. Subsequently, 1 ml GC broth was added and the samples were incubated for 5-6 h at 35°C without shaking. The cells were pelleted by centrifugation for 1 min, and resuspended in 200 µl GC broth. A 50 µl aliquot of the cell suspension was plated on CA-charcoal-kan agar and incubated for 2 days at 35°C in an atmosphere of 5% CO_2.

2.5 Construction of the Signature-tagged Transposon Vector:

2.5.1 Construction of Plasmid pBluTn:

The well characterized pBluescript (Stratagene; Cedar Creek, TX) plasmid was chosen as the vector to construct the transposon chromosomal libraries because of two significant properties. First, the plasmid possessed restriction sites that permitted the efficient cloning of the inserts and secondly, the plasmid was unable to replicate in *H. ducreyi* which facilitated the construction of the signature-tagged mutants by homologous recombination. Transposon Tn1545-Δ3 was PCR amplified from pMS1 (Table 3) using primers Tn1545K-F and Tn-1545K-R (Table 2). The PCR reaction conditions were: 2 min at 94°C, 30 cycles consisting of 30 sec at 94°C, 45 sec at 49°C and 3 min at 72°C, and a final elongation of 5 min at 72°C. Following gel purification, Tn1545-Δ3 was ligated into pBluescript that had previously been linearized with *Kpn*I using T4 DNA ligase (Invitrogen, Carlsbad, CA) according to manufacturer's instructions in a 10 µl reaction volume. The reaction mixture contained 3 µg of purified PCR product, 1 µg plasmid, 1 U T4 DNA ligase, 1 µl 10X T4 DNA ligase reaction buffer and water to a

final volume of 10 µl. Insert was added to achieve a vector: insert ratio of 1:3. Following incubation for 12 h at 4°C, the ligation mixture was introduced into *E. coli* Top10 cells by transformation. Transformants were selected on LB agar containing 50 µg/ml kanamycin and 100µg/ml ampicillin and colonies were screened by the cracking assay. Plasmid constructs were verified by restriction enzyme digestion and PCR amplification of the kanamycin cassette and of the transposon excision-integrase genes. The kanamycin cassette was amplified using R-kan and F-kan primers (Table 2) in a PCR reaction comprising 94°C for 2 min, 30 cycles consisting of 45 sec at 94°C, 30 sec at 63.7°C, and 90 sec at 72°C, and a final elongation step of 10 min at 72°C. The excision (xis) and integrase (int) genes of Tn1545-Δ3 were amplified using primers Xis-R and Int-L (Table 2) with PCR conditions as follows: 3 min initial denaturation step at 94°C, 30 cycles consisting of 45 sec at 94°C, 30 sec at 58°C, and 90 seconds at 72°C, and a final elongation step at 72°C for 10 min. The resulting plasmid construct was termed pBluTn.

2.5.2 Construction of Plasmid pBluTnSTM:

Twelve unique oligonucleotide signature tags (77) were cloned downstream of the excision and integrase genes to the transposon Tn1545-Δ3 using inverse PCR (Table 4). The signature-tags (ST) were incorporated onto the primers as linkers (Table 2). The PCR reaction used Expand Long Template enzyme mix (Roche, Boehringer Mannheim, Germany) to amplify the 6 kb plasmid product as per manufacturer's directions. The reaction conditions were as follows: initial denaturation at 92°C for 2 min, 30 cycles of 30 sec at 92°C, 45 sec at 47°C and 6 min at 68°C and the final elongation at 68°C for 10

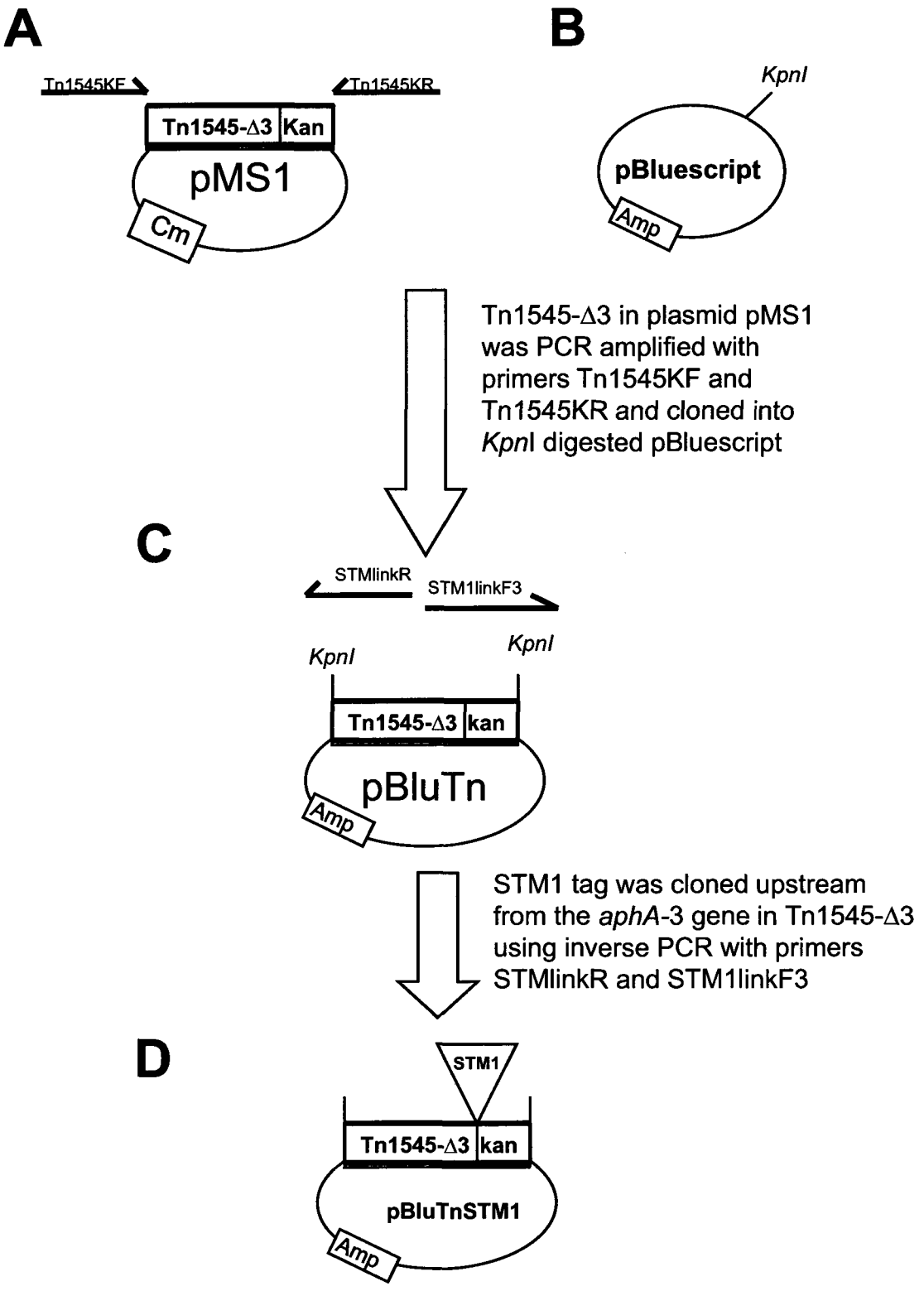
Table 4. DNA sequences of the oligonucleotides synthesized for the STM tags. Each STM tag was 21 nucleotides with a T_m = 64°C that allowed a one step PCR screening when primer combinations were used. The first 13 nucleotides of the 5'-end consist of the consensus sequences which have higher ΔGs for optimizing PCR. The red nucleotides represent the variable 3'-ends which define tag specificity and permit amplification of unique DNA sequences. Each tag was used as a primer in conjunction with a second primer synthesized within the Tn1545-Δ3 transposon (Tn1545R) in a PCR reaction.

TAG #	NUCLEOTIDE SEQUENCE
1	5'-GTACCGCGCTTAAACGTTCAG-3'
2	5'-GTACCGCGCTTAAATAGCCTG-3'
3	5'-GTACCGCGCTTAAAAGTCTCG-3'
4	5'-GTACCGCGCTTAATAACGTGG-3'
5	5'-GTACCGCGCTTAAACTGGTAG-3'
6	5'-GTACCGCGCTTAAGCATGTTG-3'
7	5'-GTACCGCGCTTAATGTAACCG-3'
8	5'-GTACCGCGCTTAAAATCTCGG-3'
9	5'-GTACCGCGCTTAATAGGCAAG-3'
10	5'-GTACCGCGCTTAATAGGCAAG-3'
11	5'-GTACCGCGCTTAACAATCGTG-3'
12	5'-GTACCGCGCTTAACTAGTAGG-3'

min. The 6 kb amplicons were gel purified and re-ligated with T4 DNA ligase for 12 h at 4°C prior to transforming *E. coli* TOP10 cells. Kanamycin and ampicillin resistant transformants were screened by colony PCR using primers specific to the STM tag (STM1-12) and the transposon (Tn1545R). The colony PCR amplification conditions were as follows: 10 min for initial denaturation at 94°C, 30 cycles consisting of 30 sec at 94°C, 45 sec at 49°C, 2 min at 72°C and a final elongation step at 72°C for 10 min. Recombinant plasmids were verified by PCR amplification using a STM specific primer (STM1-12) and Tn1545R in a PCR amplification reaction comprising a 2 min initial denaturation step at 94°C, 30 cycles consisting of 30 sec at 94°C, 45 sec at 49°C, 2 min at 72°C, and a final elongation step at 72°C for 10 min. Each STM construct in pBluTn was renamed according to the incorporated ST tag (pBluTnSTM1, pBluTnSTM2, pBluTnSTM3, pBluTnSTM4, pBluTnSTM5, pBluTnSTM6, pBluTNSTM7, pBluTnSTM8, pBluTnSTM9, pBluTnSTM10, pBluTnSTM11, and pBluTnSTM12) (Figure 2).

The specificity of the ST in each construct was determined in PCR amplification reactions using a primer to each ST and a primer to the kanamycin cassette. Each construct also underwent PCR amplification using primers F-kan and R-kan and Xis-L and Int-R to detect the presence of the kanamycin (*aphA-3*) cassette and the transposon excision and integrase genes, respectively. The constructs were confirmed by nucleotide sequencing.

Figure 2. A schematic diagram of the construction of pBluTnSTM1. The transposon Tn1545-Δ3 was cloned by PCR amplification using primers Tn1545KF and Tn1545KR from the plasmid pMS1 (A) and ligated into *Kpn*I digested pBluescript (B) creating plasmid pBluTn (C). STM1 tag was cloned 75 bp upstream from the *aphA*-3 gene in Tn1545-Δ3 using inverse PCR with primers STMlinkR and STM1linkF3 that created pBluTnSTM1 (D).


A
B
Tn1545KF
Tn1545KR
Tn1545-Δ3
Kan
pMS1
Cm
KpnI
pBluescript
Amp
Tn1545-Δ3 in plasmid pMS1 was PCR amplified with primers Tn1545KF and Tn1545KR and cloned into KpnI digested pBluescript
C
STMlinkR
STM1linkF3
KpnI
KpnI
Tn1545-Δ3
kan
pBluTn
Amp
STM1 tag was cloned upstream from the aphA-3 gene in Tn1545-Δ3 using inverse PCR with primers STMlinkR and STM1linkF3
D
STM1
Tn1545-Δ3
kan
pBluTnSTM1
Amp

2.6 Generation of a Signature-tagged Transposon Mutant Library in *H. ducreyi*:

The recombinant STM plasmids were electroporated into *H. ducreyi* 35000 as described in section 2.4.2. Transformants were screened by colony PCR using a STM tag specific primer STM1-12 and Tn1545R in a PCR reaction as described in section 2.5.2.

2.7 Growth Assays and Staining:

2.7.1 Gram's Stain:

The morphology and tinctorial properties of bacterial isolates grown on agar plates was assessed by Gram staining as previously described (30).

2.7.2 Growth Curves:

H. ducreyi HP 35000 and ST mutants grown on agar plates for 12 h at 35°C in an atmosphere of 5% CO2 were suspended in 5 ml of GC broth. The cultures were grown in the shaking incubator at 35°C with 5% CO_2 until mid-log phase was reached (OD_{600} of 0.2). A 0.5-1 ml aliquot inoculated 50-100 ml of fresh GC broth that was incubated at 35°C with 5% CO_2 inside a water-jacketed incubator on a shaker setting 4 for 24 h. The OD_{600} was measured at 2, 4, 6, 8, 10, 12 and 24 h. Aliquots were removed and serial ten-fold dilutions were plated onto CA. Each growth curve was performed three times. The mean OD_{600} values and mean viable counts (expressed as CFU/ml) were plotted against time. A Gram's stain of cells from each sampling was performed to verify the cocco-bacillary Gram negative morphology characteristic of *H. ducreyi*.

2.7.3 Competition Assays:

H. ducreyi HP 35000 and ST mutants grown on agar plates for 12 h at 35°C in an atmosphere of 5% CO2 were suspended in 5 ml of GC broth. The cultures were grown in the shaking incubator at 35°C with 5% CO_2 until mid-log phase was reached (OD_{600} of 0.2). For the competition assay, a mixture comprising 0.5-1 ml aliquots of wild type and mutant strains was inoculated 50-100 ml of fresh GC broth that was incubated at 35°C with 5% CO_2 inside a water-jacketed incubator on a shaker setting 4 for 24 h. Growth curves were conducted as described in section 2.7.2. Viable counts were determined by plating on selective (CAKan) and non-selective (CA) media as described in section 2.7.2. The agar plates were incubated at 35°C in an atmosphere of 5% CO_2 for 48 h. The competitive index (CI) was calculated relative to the wild type *H. ducreyi* 35000 and was defined as the output ratio (in CFU/ml of mutant/wild type) divided by the input ratio (in CFU/ml of mutant/wild type) during mid log phase at 8h.

2.8 Construction of a Signature-tagged *H. ducreyi* Hemoglobin A (HgbA) Mutant by Homologous Recombination:

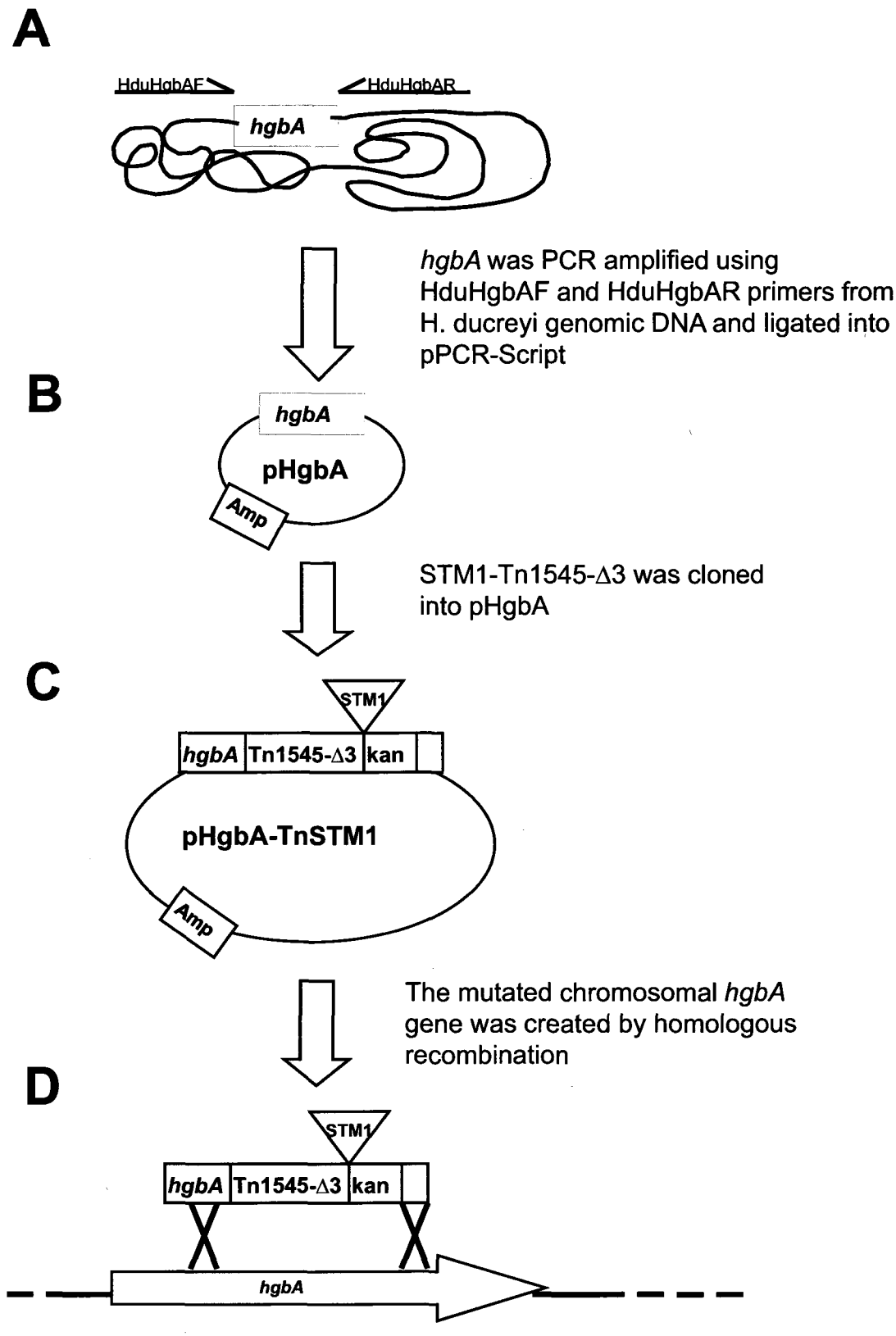
A signature-tagged *H. ducreyi* 35000 HgbA mutant was constructed to serve as a positive control for STM protocol because a previous study had demonstrated that the isogenic HgbA mutant was avirulent in the temperature-dependent rabbit model (TDRM) of chancroid infection.

The wild type *hgbA* gene was cloned by PCR amplification from the *H. ducreyi* 35000 HP genome using primers HduHgbAF and HduHgbAR (Table 2) in a PCR amplification reaction using *pfu* DNA polymerase (Stratagene; Cedar Creek, TX) under

the following parameters: initial denaturation step at 94°C for 3 min, 30 cycles consisting of 45 sec at 94°C, 30 sec at 50°C for annealing and 3 min at 72°C, and a final elongation step for 10 min at 72°C. The 3 kb PCR product was purified using the QIAgen PCR purification kit and was polished to create compatible ends for cloning into pPCR-Script (Stratagene; Cedar Creek, TX) in a reaction performed as per manufacturer's instruction. The polishing reaction consisted of 10 µl purified PCR product, 1 µl 10 mM dNTP mix, 1.3 µl 10X polishing buffer and 1 µl cloned *pfu* DNA polymerase in a final volume of 13 µl that was incubated at 72°C for 30 min. The polished PCR product was ligated into pPCR-Script as outlined per manufacturer's directions with an insert: vector ratio of 40:1. The ligation reaction contained 1 µl pPCR-Script Amp SK(+) (10 ng/µl), 1 µl pPCR-Script 10X reaction buffer, 0.5 µl 10 mM rATP, 4 µl polished PCR product, 1 µl *Srf*I restriction enzyme, 1 µl T4 DNA ligase and 1.5 µl ddH_20. The ligation reaction was incubated at RT for 1 h, heated at 65°C for 10 min and kept on ice until transformation into *E. coli* Top10 cells. Transformants were plated on LB-kan-amp-Xgal-IPTG agar and incubated overnight at 37°C. Transformants were screened by cracking assay and plasmid miniprep. The proper construct was confirmed by nucleotide sequencing. The resulting 6 kb plasmid was named pHgbA (Table 3, Figure 3).

A signature-tagged transposon Tn1545-Δ3 containing tag 1 was PCR amplified from pBluTnSTM1 using primers Tn1545K-F and Tn-1545K-R (Table 2) as described in section 2.5.1. The 3.8 kb transposon product was gel purified using QIAquick gel extraction kit. The recombinant plasmid pHgbA was digested overnight with *Avr*II in a 37°C waterbath. The linearized pHgbA was purified using the QIAquick gel extraction kit. Tn1545-Δ3 was ligated into pHgbA using T4 DNA ligase (Invitrogen, Carlsbad, CA)

Figure 3. A schematic diagram of the *hgbA* gene replacement used to create a signature-tagged *H. ducreyi hgbA* mutant. The *hgbA was* PCR amplified using HduHgbAF and HduHgbAR primers from *H. ducreyi* genomic DNA (A) and ligated into pPCR-Script to create pHgbA (B). STM1-Tn1545-Δ3 was cloned into the *Avr*II site of pHgbA to create pHgbA-TnSTM1 (C). The mutagenesis cassette containing the signature-tagged Tn1545-Δ3 transposon within the *hgbA* coding region was used to mutate the chromosomal *hgbA* gene in *H. ducreyi* 35000 by homologous recombination (D).


A
HduHgbAF
HduHgbAR
hgbA
hgbA was PCR amplified using HduHgbAF and HduHgbAR primers from H. ducreyi genomic DNA and ligated into pPCR-Script
B
hgbA
pHgbA
Amp
STM1-Tn1545-Δ3 was cloned into pHgbA
C
STM1
hgbA
Tn1545-Δ3
kan
pHgbA-TnSTM1
Amp
The mutated chromosomal hgbA gene was created by homologous recombination
D
STM1
hgbA
Tn1545-Δ3
kan
hgbA

as per the manufacturer's instructions and the construct was transformed into *E. coli* Top10 cells. Transformants were selected on LB agar containing 50 μg/ml kanamycin and 100μg/ml ampicillin. The resulting 9 kb plasmid construct was named pHgbA-TnSTM1-12 and was verified by PCR amplification using primers STM1-12 and Tn1545R in conditions described in section 2.5.2 (Figure 2).

The recombinant plasmid pHgbA-TnSTM1-12 was introduced into *H. ducreyi* by electroporation as described in section 2.4.2. Transformants were screened by colony PCR with STM1-12 and Tn1545R and positive clones bearing the mutated *hgbA* gene were verified by PCR amplification using the HduHgbAF and HduHgbAR primers. Nucleotide sequencing confirmed the presence the mutated chromosomal gene incorporating the ST.

2.9 Temperature-dependent Rabbit Model (TDRM) of Chancroid Infection:

2.9.1 Animal Strains and Husbandry:

All animal experiments were approved by the Animal Care and Veterinary Services (ACVS) staff at the University of Ottawa (Protocol BMI-71). Twenty-one age-matched male New Zealand White rabbits, between 2.5 and 3 kg in weight, were purchased from Charles River Co. (St. Constant, QC). The animals were acclimatized for one week at RT after arrival at the animal care facility. The temperature in the rabbit room was subsequently reduced by 2°C everyday to 14 ±°C one week before infection using a Thermo Air Plus air conditioning unit. Rabbits were housed individually in wire cages (Hoeltge, Inc., Cincinnati, OH) in an 11.7 m^2 room, and fed and watered *ad libitum*. In addition, each rabbit was provided with an alfalfa hay cube or a carrot, on

alternate days of the week. For environment enrichment, "bunny blocks" and "jingle balls" were provided (BioServ, Frenchtown, NJ).

2.9.2 Wellness Monitoring:

The rabbits were closely monitored on the day of the inoculation, and daily for a minimum of 7 days following. The ACVS staff were familiar with the normal health parameters of rabbits, recording the weight of the rabbits every second day. The attitude and demeanor was also noted.

2.9.3 Preparation of *H. ducreyi* Inocula:

H. ducreyi cultures were harvested at the late log phase of growth (12 h) in GC broth, pelleted by centrifugation, and then resuspended in GC broth at 10^8 CFU/ml. Ten-fold serial dilutions yielded cell suspensions at 10^7, 10^6, 10^5, and 10^4 CFU/ml. 100 μL of each suspension was drawn into a 1mL tuberculin syringe in preparation for the rabbit inoculation. An identical volume was diluted in GC broth and the dilutions were plated onto CA to determine the viable delivered inoculum.

2.9.4 STM Mutant Pools for TDRM Inoculation:

Individual STM mutants were selected on the basis of their growth curves compared to wild type and their *in vitro* competitive index (CI). *H. ducreyi* cultures were harvested at the late log phase of growth in GC broth, pelleted by centrifugation, and then resuspended in GC broth at 10^9 CFU/ml. Ten-fold serial dilutions yielded cell suspensions at 10^8, 10^7, 10^6, 10^5, and 10^4 CFU/ml. For the STM pools, 1 ml from each of

the bacterial solutions was pooled together and mixed (Figure 4). Bacterial solutions were kept at 35°C prior to inoculation of 100 µl intraepithelially on a shaved rabbit back.

2.9.5 ST-HgbA Mutant:

The ST-HgbA mutant was included in one of the STM pools that were inoculated into the TDRM. It was grown from frozen stock and an overnight culture was started as described in section 2.2. Ten-fold serial dilutions were made and the appropriate dilution was pooled together with the other STM mutants.

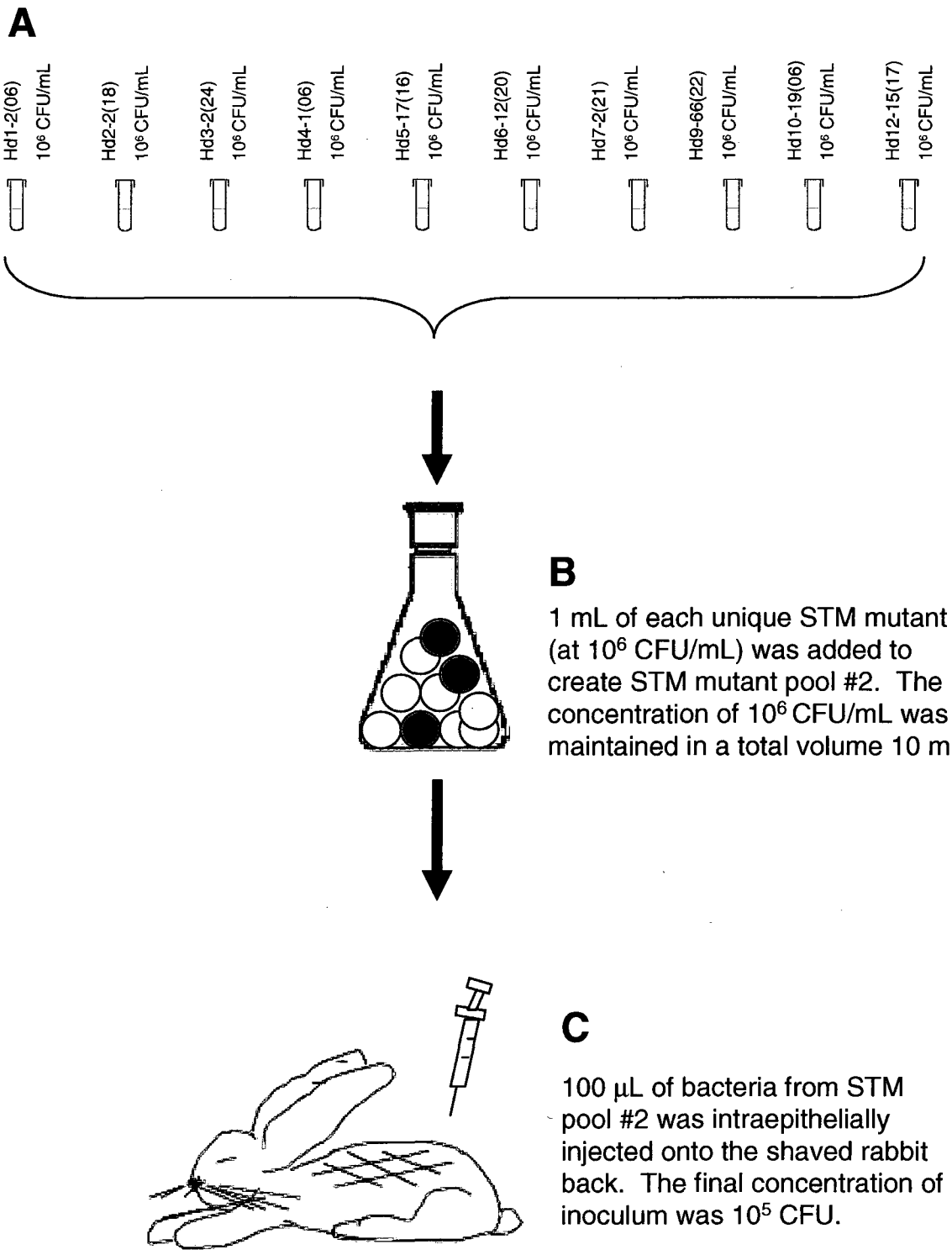
2.9.6 Determination of the Optimal Inoculum Dose:

The temperature dependent rabbit model (TDRM) of *H. ducreyi* infection has been described in detail (38, 53). Rabbits were challenged with *H. ducreyi* as described by Desjardins et al 1995 (37). It was crucial that the ambient room temperature was maintained at 14°C for the duration of the experiment. A ducted fan was used to eliminate temperature gradients as well as limited and minimal access to the rabbit room. The rabbit backs were shaved 1 day before inoculation with *H. ducreyi* or the STM mutants and every day thereafter. To determine the optimal inoculum dose, five triplicate inocula (10^8, 10^7, 10^6, 10^5, and 10^4 CFU) were injected in 100µL volumes intraepithelially in a 3x5 grid drawn on each rabbit back.

2.9.7 Extraction and Recovery of *H. ducreyi* from the Rabbit Lesions:

At day 4, full presentation of the ulcer was usually observed in the rabbit and by day 5 it was observed that some of the lesions would begin to resolve. Accordingly, day

Figure 4. A Schematic diagram of STM pool #2 inoculum size for the TDRM. (A) Each STM mutant was adjusted to a concentration of 10^6 CFU/mL. (B) 1 mL of each unique STM mutant (at 10^6 CFU/mL) was added to create STM mutant pool #2. The concentration of 10^6 CFU/mL was maintained in a total volume 10 mL. (C) 100 μL of bacteria from STM pool #2 was intraepithelially injected onto the shaved rabbit back. The final concentration of inoculum was 10^5 CFU.


A
Hd1-2(06) 10⁶ CFU/mL
Hd2-2(18) 10⁶ CFU/mL
Hd3-2(24) 10⁶ CFU/mL
Hd4-1(06) 10⁶ CFU/mL
Hd5-17(16) 10⁶ CFU/mL
Hd6-12(20) 10⁶ CFU/mL
Hd7-2(21) 10⁶ CFU/mL
Hd9-66(22) 10⁶ CFU/mL
Hd10-19(06) 10⁶ CFU/mL
Hd12-15(17) 10⁶ CFU/mL
B
1 mL of each unique STM mutant (at 10⁶ CFU/mL) was added to create STM mutant pool #2. The concentration of 10⁶ CFU/mL was maintained in a total volume 10 mL.
C
100 µL of bacteria from STM pool #2 was intraepithelially injected onto the shaved rabbit back. The final concentration of inoculum was 10⁵ CFU.

4 of the experiment was chosen as the endpoint to ensure full recovery of all STM mutants inoculated before any mutants died off. Prior to excision, each rabbit was administered an anesthetic and euthanized as per ACVS standard operating procedure. The entire lesion at each inoculum was excised by ACVS staff. Removal of the lesion was performed by lifting the skin away from the muscle with tweezers and excising the lesion from the skin. The lesion was placed in a sterile 10 ml specimen jar containing 2 ml of 1X sterile PBS before processing. Excess fur surrounding the excised lesion was trimmed with scissors and then the sample was cut up into smaller pieces. The tissue was homogenized using an Omni 2000 homogenizer (Omni International, Waterbury CT) for 2 min in 2 ml 1X PBS. 100 μL of each suspension was diluted in 1X PBS and the dilutions were plated onto CA, CA-charcoal-kan (30 μg/ml) or CA-charcoal-kan (30 μg/ml)-van (3 μg/ml) agar to recover and quantitate the bacteria in the rabbit lesions.

The recovered bacteria were stored at -80°C and genomic DNA from the mixed sample of mutants was extracted as described in section 2.3.8. The genomic DNA was verified by electrophoresis on an agarose gel, quantified using a Genequant II DNA calculator or by comparison of band intensity to a commercially available High Mass ladder and then subjected to STM tag screening by PCR. The PCR was performed using primers STM1-12 and Tn1545R as described in section 2.5.2. These PCR samples were referred to as the output pool.

2.9.8 Histopathology of the Rabbit Lesions:

Excised lesions were submerged in 5 ml of 10% (v/v) formalin for 24 h. Samples were embedded in paraffin, and tissue sections were transferred onto clean microscope

slides, followed by staining with hematoxylin and eosin (H&E), according to the standard processing protocol established by the Department of Pathology at the University of Ottawa. The samples were visualized using a Nikon Imaging Microscope Eclipse 80i (Melville, NY) and the images captured with a digital camera. The interpretation of the histopathology of the tissue sections was performed in collaboration with Dr. Susan Robertson (Anatomical Pathology, Ottawa Hospital)

2.9.9 STM-TDRM:

The rabbits' backs were shaved 1 day before inoculation with *H. ducreyi* or the STM mutants and every day thereafter. The inocula were injected, in 100 µL volumes intraepithelially in a 3x5 grid drawn on each rabbit back. Three different STM pools of individually tagged *H. ducreyi* mutants were tested in each rabbit. Each STM pool was inoculated at 10^5, and 10^4 CFU in duplicate on the rabbit back. The positive control comprised the wild type *H. ducreyi* 35000 strain inoculated at 10^4 CFU in duplicate on the rabbit back. The negative control consisted of a 100 µl injection from a 5 ml *H. ducreyi* culture heated at 100°C for 10 min or 100 µl of spent GC broth. These two controls were used to test for any inflammatory response to the samples injected. One negative control was injected on each rabbit back. Two rabbits were used in each experiment.

The rabbit backs were shaved and observed daily. The lesions at each inoculum were assigned a clinical score (1= redness, 2= induration, 3= suppuration, 4= ulceration). The lesions were excised at day 4 as described in section 2.9.7.

2.9.10 Individual STM-mutant Virulence Attenuation in TDRM:

Candidate attenuated STM mutant strains were inoculated individually into the TDRM to reconfirm the observed avirulence in the first screen when they were tested within a pool of STM mutants. Five triplicate inocula at 10^4 CFU were injected, in 100 μL volumes, intraepithelially in a 3x5 grid drawn on each rabbit's back.

Two lesions at each inoculum were measured using electronic calipers, and assigned a clinical score (1= redness, 2= induration, 3= suppuration, 4= ulceration). Statistical analysis was performed using a Student's t-test to compare wild type strain and mutant strain. The third lesion was cultured by injection of 100μL PBS into the lesion followed by aspiration of PBS, blood and inflammatory exudates from the lesion. This aspirate was plated on CA, and colonies were counted 2 days after plating. Culture positivity was determined by the presence of one or more colonies characteristic of *H. ducreyi* (small, gray colonies that can be pushed intact across the plate with the point of a wire probe).

CHAPTER 3: RESULTS

3.1 Construction of a Transposon Vector:

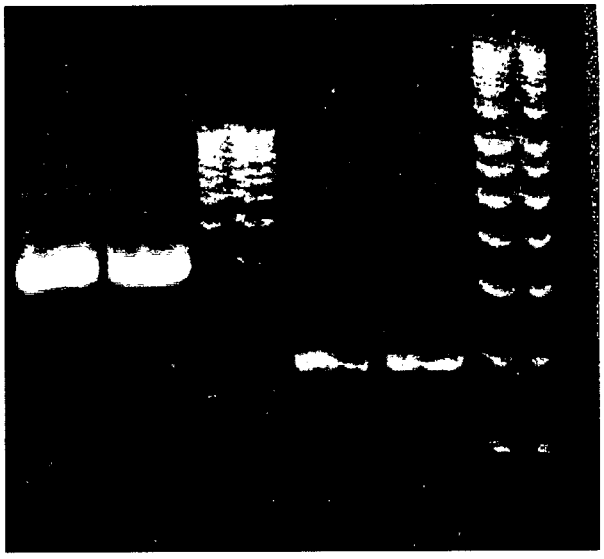
3.1.1 Cloning Transposon Tn1545-Δ3 into pBluescript II KS+:

The plasmid pBluescript was used as the backbone vector because it was better characterized vector for cloning with a complete sequenced restriction map. It also served as a suicide vector because it is unable to replicate in *H. ducreyi* and thus for survival purposes the transposon would have to integrate into the bacterial genome. The cloning vector pBluescript was *Kpn*I digested to linearize the plasmid in preparation for cloning the transposon into it (Figure 5A). Transposon Tn1545-Δ3 was PCR amplified from pMS1 using primers Tn1545K-F and Tn-1545K-R. The 3.7 kb transposon product (Figure 5B) was gel purified using QIAquick gel extraction kit to remove any impurities. Tn1545-Δ3 was ligated into pBluescript and transformed into *E. coli* Top10 cells on LB kan(50 μg/ml)-Ap(100μg/ml). The resulting 6Kb plasmid construct was name pBluTn.

The construct pBluTn was first screened by a cracking assay and transformants with the correct plasmid size of 6.7 kb were isolated and grown overnight in LB broth for miniprep (Figure 6A). The minipreps were screened by digestion with *Pst*I which cuts Tn1545-Δ3 twice resulting in a 4.5 kb and a 2.2 kb fragment. A sequential digest was performed using *Xmn*I which cuts the pBluescript backbone resulting in three fragments of sizes 1 kb, 2.2 kb, and 3.5 kb (Figure 6A). The pBluTn construct was screened by PCR using F-kan and R-kan primers to amplify the kanamycin cassette in Tn1545-Δ3 giving a 1 kb product (Figure 6B). PCR with Xis-R and Int-L to amplify the excision and integrase genes from Tn1545-Δ3 resulted in a 0.2 kb amplicon (Figure 6C).

Figure 5. Agarose gels of pBluTn cloning. (A) Agarose gel of pBluescript plasmid miniprep and pBluescript *Kpn*I digested (3 kb) that linearized the vector for cloning Tn1545-Δ3 into it. Lanes 1 and 2; pBluescript II KS (+) *Kpn*I digested. Lanes 4 and 5 were pBluescript minipreps. The sizes of the standards, lane 3 (1 kb ladder) and lane 6 (supercoiled ladder), in kilobases were shown. (B) PCR amplification of Tn1545-Δ3 using primers Tn1545K-F and Tn1545K-R from pMS1 that gave a 3.8 kb product for cloning into pBluescript used to create pBluTn. Lanes 2-5, Tn1545-Δ3 amplicon. Lane 1 showed the 1 kb ladder with standard sizes in kilobases.

A


1
2
3
4
5
6
4075
3054
2036
1636
6070
5012
3990
2972
2067

B

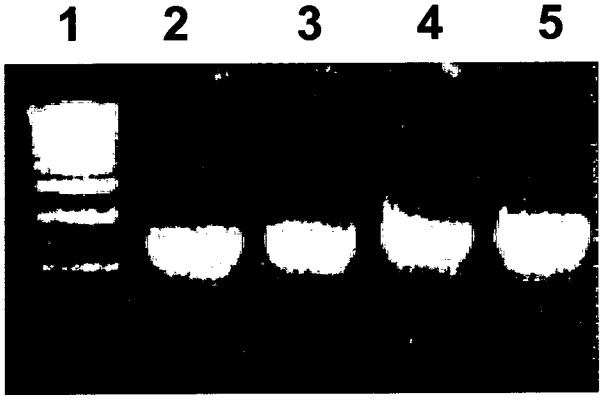
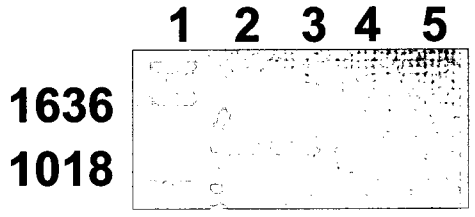

1
2
3
4
5
4075
3054

Figure 6. Agarose gels to verify pBluTn vector. (A) pBluTn miniprep and restriction enzyme digest to assess the size of construct. Lane 2 was the 6 kb plasmid pBluTn and lane 4 was pBluTn digested with *Pst*I and *Xmn*I that resulted in 3 fragments. Lane 1 was a supercoiled ladder and lane 3 was a ReadyLoad 1 kb ladder. Sizes of the standards were shown in kilobases. (B) PCR amplification of the kanamycin cassette from Tn1545-Δ3 using primers F-kan and R-kan. Lane 1 was the 1 kb standard. Lane 2 was the positive control pMS1 and lane 5 was a negative PCR control of dH_20. Lanes 3 and 4, templates were pBluTn constructs. (C) PCR amplification of the excision and integrase genes of Tn1545-Δ3 using primers Xis-R and Int-L. Lane 1, pMS1 positive control and lane 4 was a PCR negative control of dH_2O. Lanes 2 and 3 were pBluTn constructs. Lane 5 was 1 kb standard.

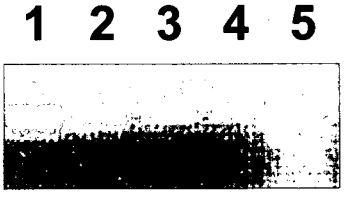
A


1 2 3 4
6030
5012
3990
2972
2067
4072
3054
2036
1636
1018

B


1 2 3 4 5
1636
1018

C


1 2 3 4 5
200

3.1.2 Transposition of pBluTn in *H. ducreyi*:

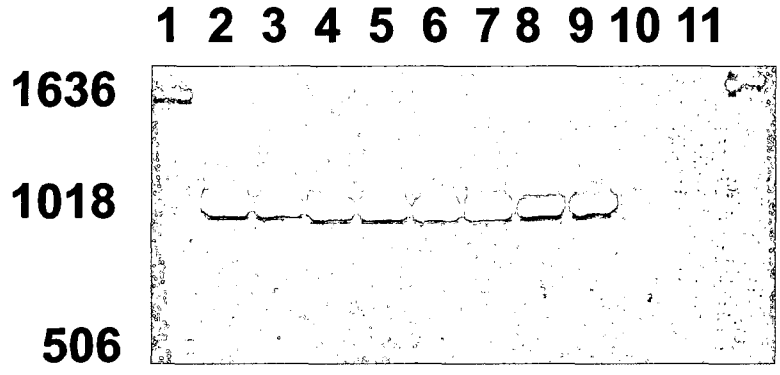
The pBluTn plasmid was electroporated in *H. ducreyi* to verify that the transposon construct was functional. Transformants were selected on CA-charcoal-Kan (30μg/ml) plates and counter selected on CA-charcoal-Ap (2μg/ml). The transformants were screened by colony PCR using F-kan and R-kan primers that produced a 1 kb band and also using Xis-R and Int-L primers that produced a 0.2 kb fragment (Figure 7A and B). For the electroporation, pLS88 was used as a positive control because it was a native plasmid of *H. ducreyi*. Transformation with pLS88 also resulted in kanamycin resistant colonies but did not amplify the 1 kb band using primers F-kan and R-kan because of the different kanamycin cassette (Figure 7A). Stocks of the mutants were made with freezing media and stored at -80°C.

The stability of the transposon in *H. ducreyi* transformants was assessed and confirmed by serial passage of the mutants every 2 days on selective media CA-charcoal-kan (30μg/ml) to non-selective media CA for 8 weeks. Transposon mutants were still viable at the end of 8 weeks. Colony PCR with F-kan and R-kan primers that produced a 1 kb band and, Xis-L and Int-R that showed a 0.2 kb fragment confirmed the presence of the transposon.

Random incorporation of Tn1545-Δ3 into the *H. ducreyi* genome was assessed by Southern blot. *H. ducreyi* mutant genomic DNA was extracted, digested (with *Sal*I, *Cla*I and *Aval*) and then electrophoresed on a 1% agarose gel. These restriction enzymes do not cut within the 1 kb kanamycin (*aphA3*) cassette probe. After transfer onto a nylon membrane, the blots were hybridized with a 1 kb probe specific to the Tn1545-Δ3 kanamycin cassette. Figure 8 shows the Southern analysis of 11 different *H. ducreyi*

Figure 7. Agarose gels of *H. ducreyi* transformant colony PCR after electroporation with pBluTn to verify transposition of Tn1545-Δ3. (A) Colony PCR amplification of kanamycin cassette from Tn1545-Δ3 using primers R-kan and F-kan. Lanes 2-9, positive *H. ducreyi* transformants. Lane 10 was a negative control for transposition using pLS88 and lane 11 was a negative control for PCR using dH_2O as template. Lane 1 was a 1 kb standard. (B) Colony PCR amplification of the excision and integrase genes of Tn1545-Δ3 using Xis-R and Int-L. Lanes 2 and 3 were *H. ducreyi* transformants. Lane 4 was a negative dH_2O control and lane 1 was a kb standard.

A



B

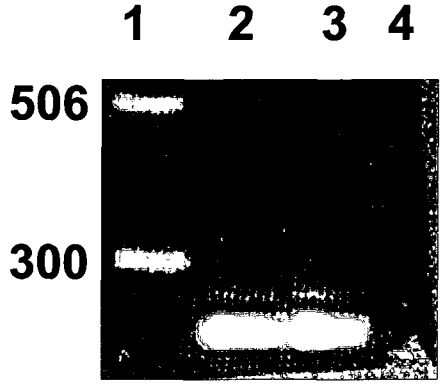
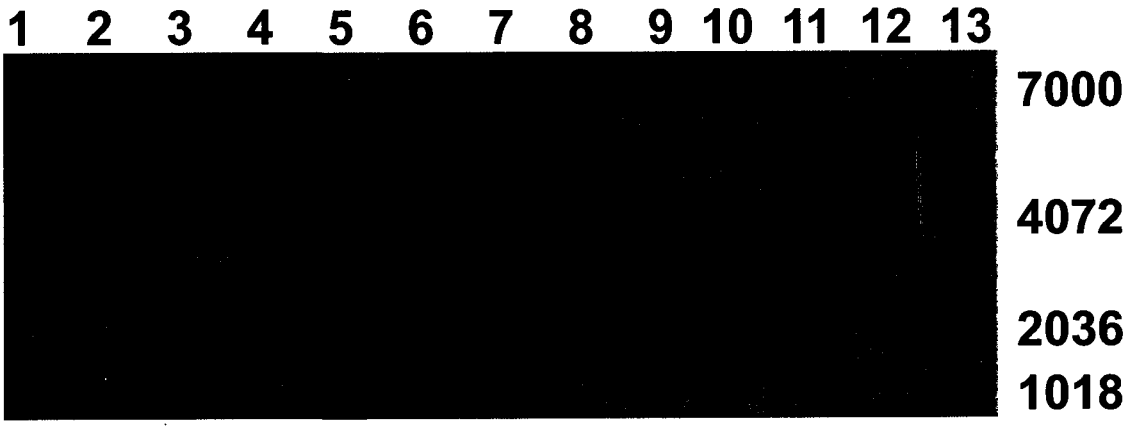


Figure 8. Southern blot analysis of *H. ducreyi* mutants. Genomic DNA obtained from clones was digested with *Sal*I, *Cla*I and *Ava*I. Ten micrograms of digested DNA was applied to a 0.6% agarose gel, electrophoresed, blotted on nylon filter, and subjected to Southern hybridization analysis using the *aphA3* gene as a probe. Lanes 1- 12, were *H. ducreyi* transposon clones HdTn1545-4, 6, 9, 11, 15, 21, 25, 31, 32, 37 and 40. Lane 13 was wild type untransformed *H. ducreyi* 35 000. Lane 12 was pMS1 to serve as a positive control. The positions of the size markers were shown in kilobases on the right.


1 2 3 4 5 6 7 8 9 10 11 12 13
7000
4072
2036
1018

transposon mutants, HdTn1545-4, 6, 9, 11, 15, 21, 25, 31, 32, 37 and 40. The incorporation of the transposon was random based on the probe annealing to different areas on the blot and thus different areas of the genome (Figure 8). A few of the mutants did have multiple copies of the transposon and were not used in the subsequent experiments.

Growth curves of Tn1545-Δ3 *H. ducreyi* mutants were also performed and compared to wild type (Figure 9). The mutants HdTn1545-4, 15, 21, and 37 displayed growth characteristic of wild type.

3.2 Construction of a Signature-tagged Transposon:

3.2.1 Cloning of a Unique Signature-tag Adjacent to Tn1545-Δ3:

The 12 unique signature tags (77) were cloned adjacent to the transposon Tn1545-Δ3 using inverse PCR. The STM oligonucleotides that were synthesized as tags, were each 21 nucleotides long with a T_m of 64°C that allows for one step PCRs when primer combinations are used for screening. The consensus 5'-ends are comprised of the first 13 nucleotides and have higher ΔGs for optimizing PCRs. The variable 3'-ends indicated in red define tag specificity and permit amplification of specific DNA fragments. Each tag is used as a primer in PCR with a primer synthesized from the Tn1545-Δ3 transposon (Tn1545R). The signature-tags, Table 4, were placed on the primers as linkers (Table 2). After successful PCR, the 6.7 kb amplicons were gel purified, re-ligated overnight and transformed into *E. coli* TOP10 cells (Figure 10A). Kanamycin and ampicillin resistant colonies were first screened by colony PCR screened using primers specific to the STM tag (STM1-12) and the transposon (Tn1545R) that produced a 2 kb band (Figure9B).

Figure 9. Growth assay of *H. ducreyi* Tn1545 transformants HdTn1545-4, 15, 21, and 37 cultures compared to wild type *H. ducreyi* 35000 for 24 h at 35°C with 5% CO_2. Aliquots were taken every 2 h for 12 h. Each value represented the mean of three triplicate cultures. Standard deviations of the mean were shown by error bars. Absence of error bars indicated that the standard deviation of the three replicate cultures was too small to be represented on the graph.

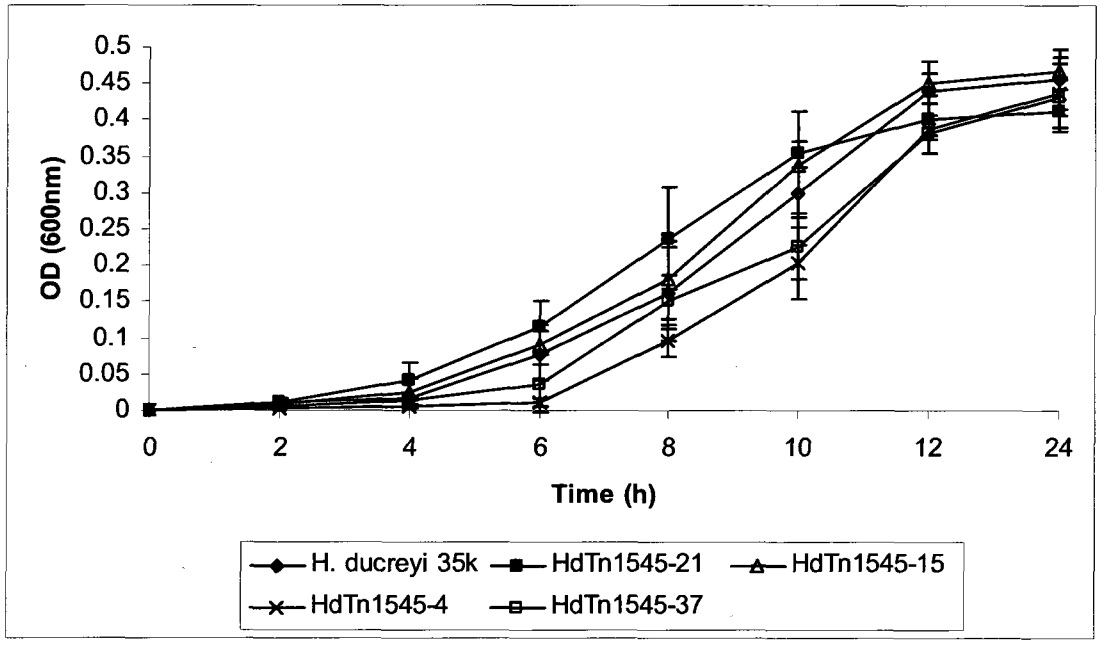
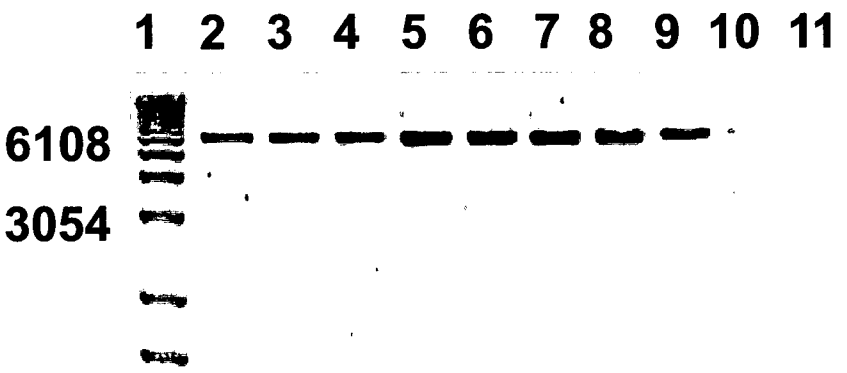
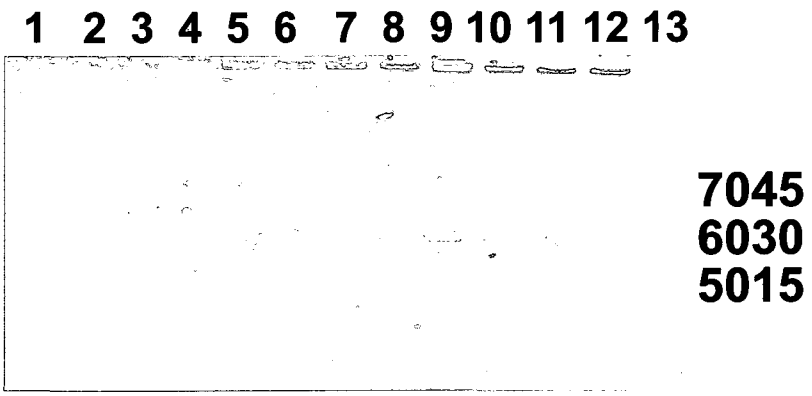

0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
OD (600nm)
0
2
4
6
8
10
12
24
Time (h)
H. ducreyi 35k
HdTn1545-21
HdTn1545-15
HdTn1545-4
HdTn1545-37

Figure 10. Agarose gels of pBluTnSTM cloning. (A) Inverse PCR using STMlink primers on pBluTn was performed to clone STM tags downstream of Tn1545-Δ3. Lanes 2 – 10 show pBluTnSTM1-9, respectively. Lane 11 was a negative PCR control with no template. Lane 1 shows the standard sizes in kilobases on the left. (B) Cracking assay of *E. coli* pBluTnSTM transformants. Lanes 1-12 were pBluTnSTM1-12, respectively. Lane 13 was a supercoiled ladder. Sizes were in kilobases illustrated on the right.

A


1 2 3 4 5 6 7 8 9 10 11
6108
3054

B


1 2 3 4 5 6 7 8 9 10 11 12 13
7045
6030
5015

STM tag positive colonies were then screened by cracking assay (Figure 10B) and transformants with the correct 6.7 kb size were picked for minipreps as seen in Figure 10A. The pBluTnSTM constructs were digested with *Pst*I and *Xmn*I which resulted in 3 fragments of sizes 1 kb, 2.2 kb and 3.5 kb as seen in Figure 11B. Each pBluTnSTM construct was also PCR screened with F-kan and R-kan (1 kb band) and Xis-L and Int-R (0.2 kb band) that amplified regions on the transposon (Figure 12A and B).

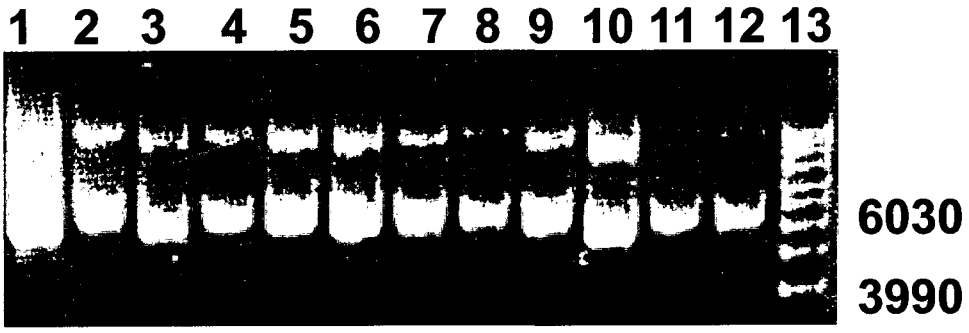
The plasmids extracted from the Top10 cells were also PCR screened using a STM specific primer (STM1-12) and Tn1545R. Figures 12 and 13 demonstrated the specificity of the tags (STM1-12 primer) in each of the pBluTn constructs that produced a 2 kb amplicon. Each STM construct in pBluTn was renamed according to the ST tag that was cloned into it (pBluTnSTM1, pBluTnSTM2, pBluTnSTM3, pBluTnSTM4, pBluTnSTM5, pBluTnSTM6, pBluTNSTM7, pBluTnSTM8, pBluTnSTM9, pBluTnSTM10, pBluTnSTM11, and pBluTnSTM12) (Figure 2). Figures 12 and 13 demonstrate specificity of the ST because each construct was tested with each STM primer in a multiplex PCR and only the STM primer corresponding to the ST that was cloned into the vector displayed a band around 2 kb. Each construct was also verified by sequencing.

3.2.2 Transposition of pBluTnSTM in *H. ducreyi*:

The signature-tagged transposon Tn1545-Δ3 from pBluTnSTM was electroporated into *H. ducreyi* as previously described. Kanamycin resistant colonies grown on CA-charcoal-Kan (30 μg/ml) were screened by colony PCR using a STM specific primer (STM1-12) and a transposon Tn1545-Δ3 specific primer (Tn1545R)

Figure 11. Agarose gels of pBluTnSTM1-12 minipreps and restriction enzyme digests. (A) Lanes 1-12 were pBluTnSTM1-12 minipreps, respectively. Lane 13 was the supercoiled ladder. The sizes of the standard in kilobases are on the right. (B) Plasmid pBluTnSTM1-4 digested with *Pst*I and *Xmn*I. Lanes 2-5 were undigested pBluTnSTM1, 2, 3, and 4. Lanes 7-10 were *Pst*I and *Xmn*I digested pBluTnSTM 1, 2, 3, and 4. Lane 1 was the supercoiled ladder and lane 6 was the 1 kb ladder. Sizes of the standard were in kilobases illustrated on the right and left.

A


1 2 3 4 5 6 7 8 9 10 11 12 13
6030
3990

B

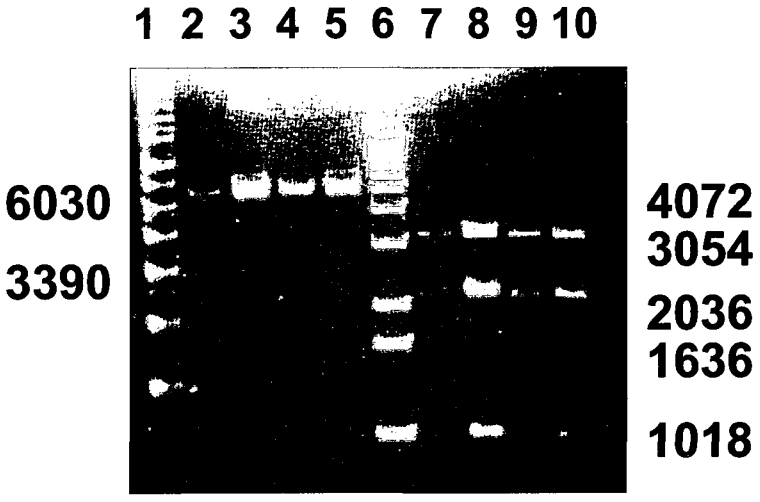
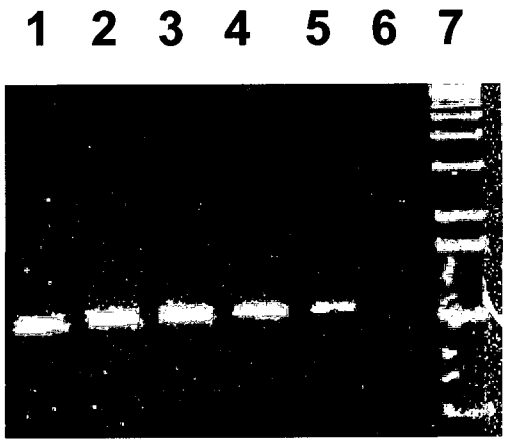

1 2 3 4 5 6 7 8 9 10
6030
3390
4072
3054
2036
1636
1018

Figure 12. Agarose gels of pBluTnSTM PCR to verify the presence of transposon Tn1545-Δ3. (A) PCR amplification of the kanamycin cassette in pBluTnSTM1-4 using primers R-kan and F-kan. Lanes 1-4 were of pBluTnSTM1, 2, 3, and 4. Lane 5, was pMS1 which was used a positive control and lane 6, was a PCR negative control of dH_2O. Lane 7 was the 1 kb ladder. (B) PCR amplification of the excision and integrase genes in pBluTnSTM1-4 using primers Xis-R and Int-L. Lanes 2-5, were pBluTnSTM1, 2, 3 and 4. Lane 5, was pMS1 which was used a positive control and lane 1, was a PCR negative control of dH_2O. Lane 7 was the 1 kb ladder. Sizes of the standards were in kilobases.

A



B

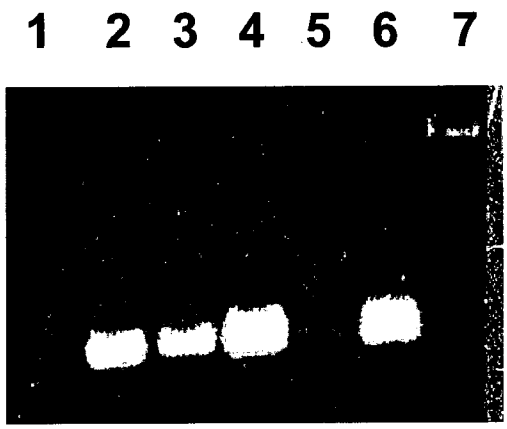


Figure 13. Multiplex PCR amplification of pBluTnSTM1-12 vectors using STM1-6 primers and Tn1545R to verify STM tag specificity. Each pBluTnSTM1-12 construct was tested with all the STM primers. (A) STM1 (B) STM2 (C) STM3 (D) STM4 (E) STM5 (F) STM6. Sizes were in kilobases depicted in between the gels.

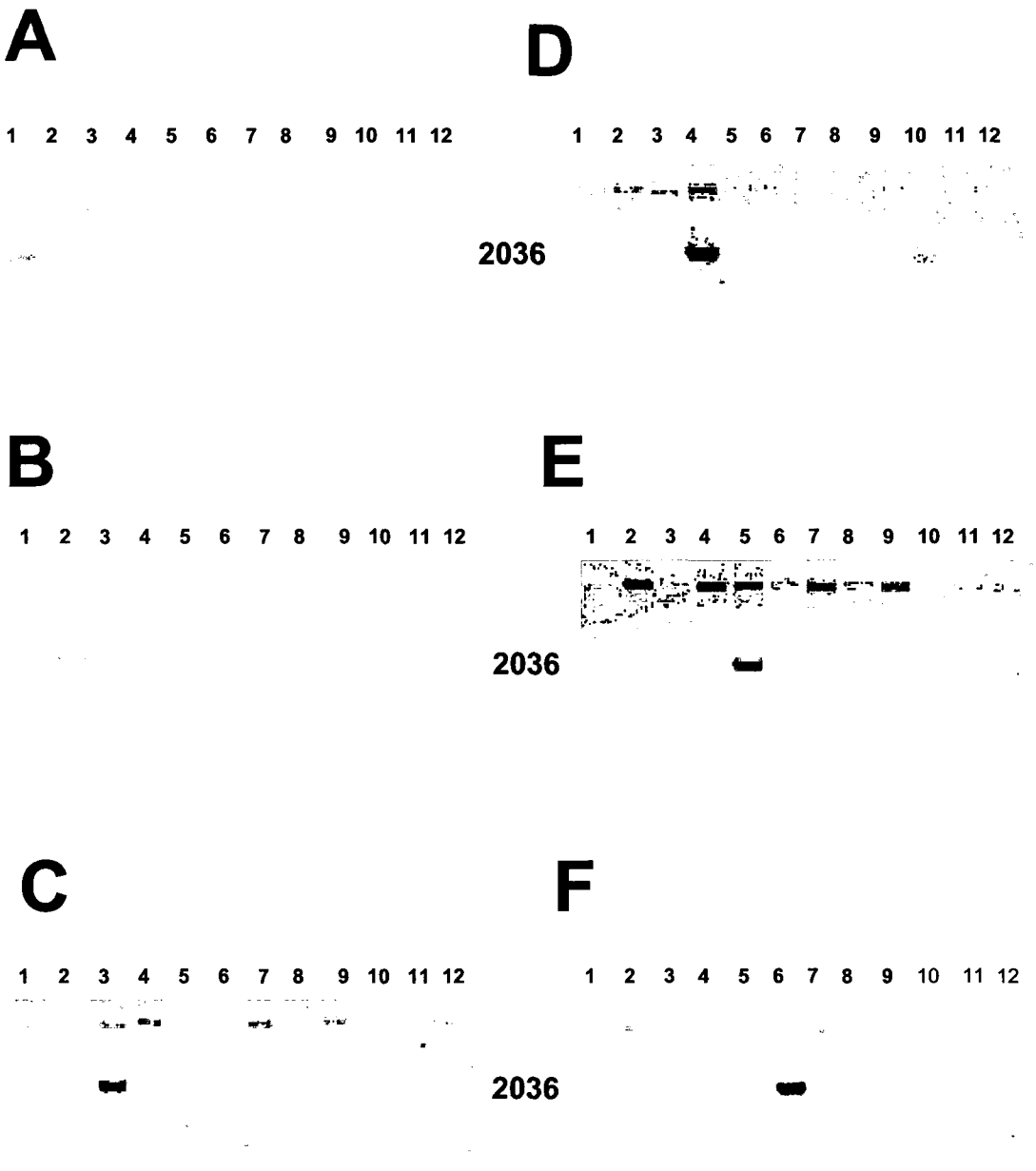
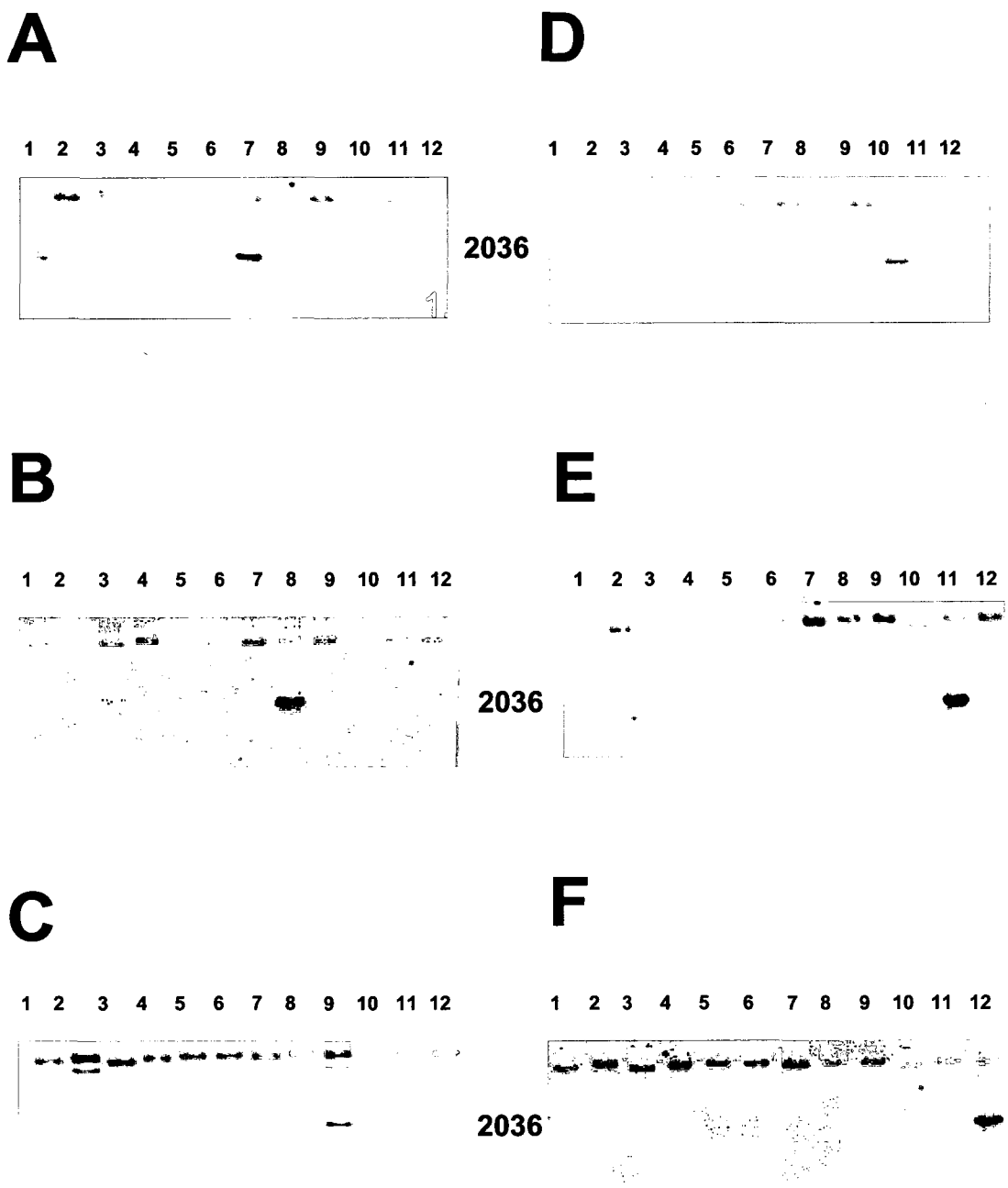

A
1 2 3 4 5 6 7 8 9 10 11 12
D
1 2 3 4 5 6 7 8 9 10 11 12
2036
B
1 2 3 4 5 6 7 8 9 10 11 12
E
1 2 3 4 5 6 7 8 9 10 11 12
2036
C
1 2 3 4 5 6 7 8 9 10 11 12
F
1 2 3 4 5 6 7 8 9 10 11 12
2036

Figure 14. Multiplex PCR amplification of pBluTnSTM1-12 vectors using STM7-12 primers and Tn1545R to verify STM tag specificity. Each pBluTnSTM1-12 construct was tested with all the STM primers. (A) STM7 (B) STM8 (C) STM9 (D) STM10 (E) STM11 (F) STM12. Sizes were in kilobases illustrated in between the gels.


A
1 2 3 4 5 6 7 8 9 10 11 12
2036
D
1 2 3 4 5 6 7 8 9 10 11 12
B
1 2 3 4 5 6 7 8 9 10 11 12
2036
E
1 2 3 4 5 6 7 8 9 10 11 12
C
1 2 3 4 5 6 7 8 9 10 11 12
2036
F
1 2 3 4 5 6 7 8 9 10 11 12

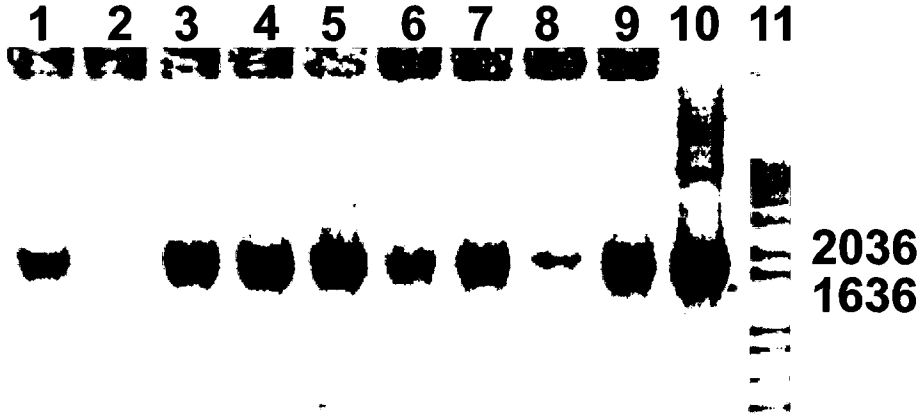
(Figure 15A). No transformants for STM11 were obtained and thus were not used for STM. *H. ducreyi* STM transformants were counterselected on CA-charcoal-Ap (2 μg/ml) plates to screen out the pBluescript backbone (DNS). Plasmid minipreps were also performed and pBluTn plasmids were not recovered (DNS). Gram stains of transformants were done to verify identity of *H. ducreyi* STM mutants. The Gram stains showed morphology typical of Gram negative *H. ducreyi* cocco bacilli (DNS). *H. ducreyi* STM mutant colonies could also be pushed around intact on agar plates and were not viable on Blood agar plates, both phenotypes are characteristic of wild type *H. ducreyi* 35000.

H. ducreyi STM mutants were serially passaged on selective CA-charcoal-Kan(30 μg/ml) and on non-selective CA-charcoal to confirm the stability of transposon. After 8 weeks, mutants were still viable and contained the transposon as assessed by colony PCR with F-kan and R-kan primers (Figure 15B).

Growth curves of ST Tn1545-Δ3 *H. ducreyi* mutants were also performed and compared to wild type (Figure 17). *H. ducreyi* STM mutants displayed similar growth characteristics as wild type, with an initial lag phase, log phase, and stationary phase (OD_{600}). The final OD_{600} for *H. ducreyi* after 24 h never exceeded 0.5. Anything greater than an OD= 0.5 meant that the culture was contaminated which was verified by Gram stain.

Figure 15. *H. ducreyi* Tn1545-Δ3 STM7 transformants constructed by electroporation of pBluTnSTM7 into *H. ducreyi* 35000. The stability of the transposon was verified in the mutants by serial passage for 8 weeks and PCR. (A) Colony PCR screening of HSTM7 mutants using STM7 and Tn1545R primers. Lanes 1-9 were HSTM7-1, 2, 3, 4, 5, 6, 7, 8, 9 mutants. Lane 2, was of HSTM7-2 did not contain the STM transposon. All other lanes show a 2 kb product. Lane 10 was the positive control for the PCR reaction containing pBluTnSTM7 as template. Lane 11 was the 1 kb ladder. (B) Colony PCR amplification of the kanamycin cassette from Tn1545-Δ3 using primers R-kan and F-kan of the HSTM7 mutants after 8 weeks of subculturing on agar. Lanes 2-11 were of HSTM7-1, 3, 4, 5, 6, 7, 8, 9, 10 and 11, respectively. Lane 2 and 12 was the 1 kb ladder. Sizes of the standard were in kilobases shown on the right.

A



B

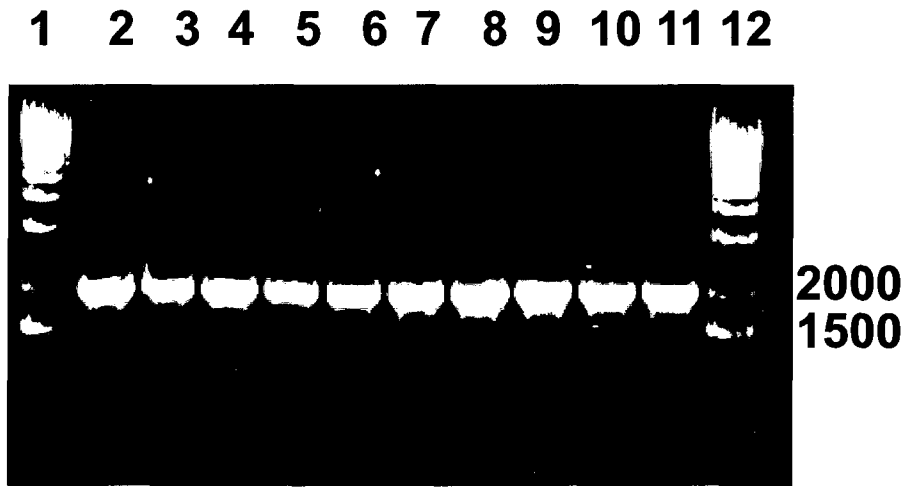
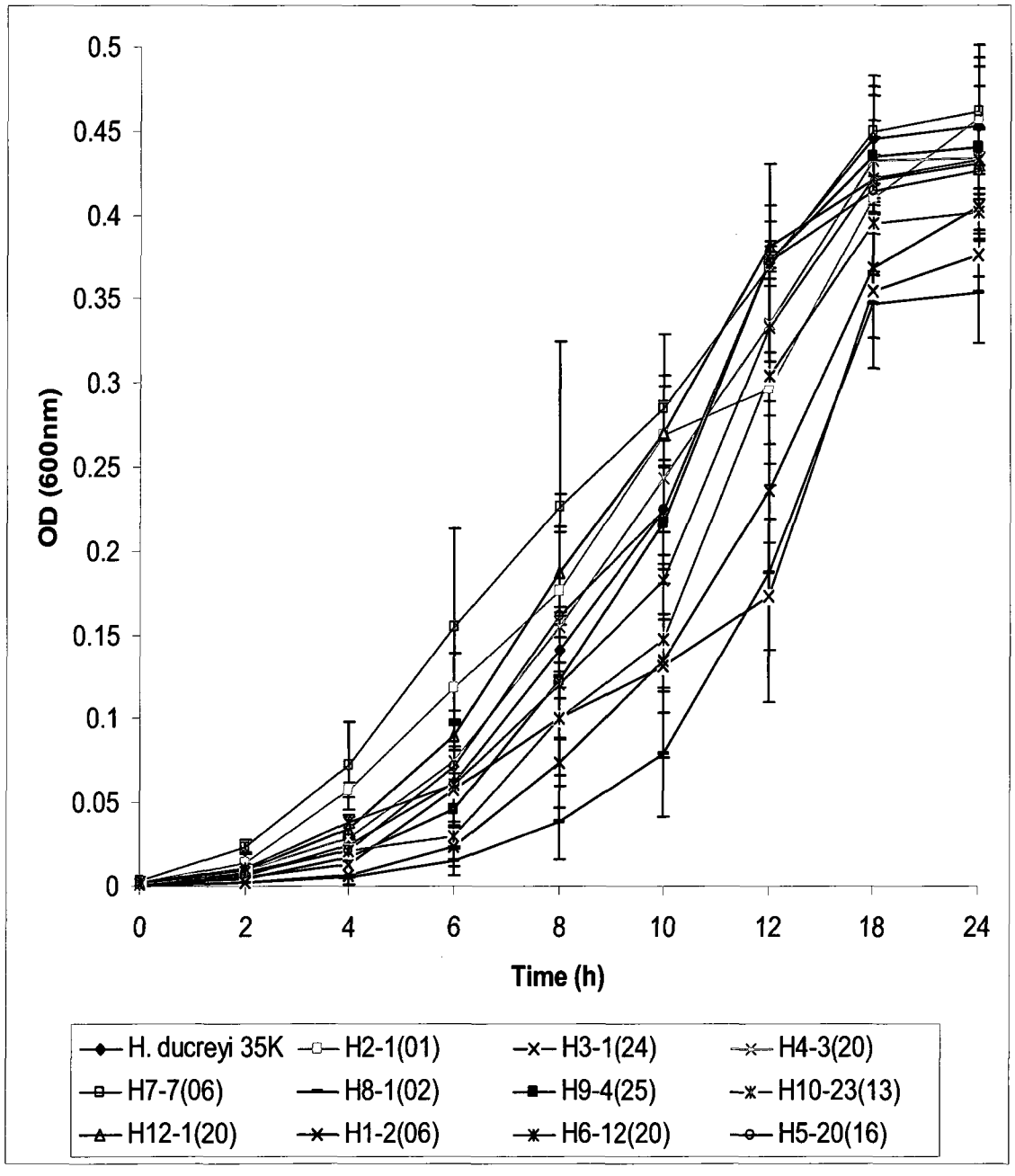


Figure 16. *H. ducreyi* STM mutant growth assay of bacterial cultures grown for 24 h at 35°C with 5% CO_2. Aliquots were taken every 2 h for 12 h and the OD_{600} was measured. The STM mutants that were examined were H1-2(06), H2-1(01), H3-1(24), H4-3(20), H5-20(16), H7-7(06), H8-1(02), H9-4(25), H10-23(13) and H12-1(20). *H. ducreyi* 35000 was used as a control. Each value represented the mean of three triplicate cultures. Standard deviations of the mean were shown by error bars. Absence of error bars indicated that the standard deviation of the three replicate cultures was too small to be represented on the graph.


0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
OD (600nm)
0
2
4
6
8
10
12
18
24
Time (h)
H. ducreyi 35K
H2-1(01)
H3-1(24)
H4-3(20)
H7-7(06)
H8-1(02)
H9-4(25)
H10-23(13)
H12-1(20)
H1-2(06)
H6-12(20)
H5-20(16)

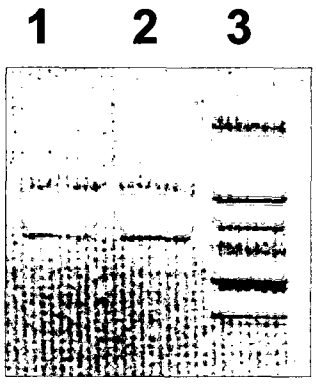
3.3 Construction of the Signature-tagged Hemoglobin A (HgbA) Mutant Via Homologous Recombination:

An isogenic HgbA mutant has been previously characterized to be avirulent in the TDRM of chancroid infection (135). A signature-tagged HgbA mutant would serve as a positive control for STM.

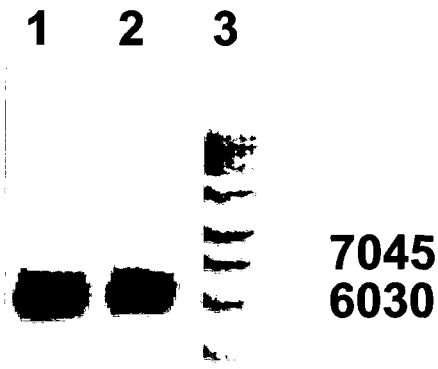
The *hgbA* gene was PCR amplified from *H. ducreyi* 35000 HP using primers HduHgbAF and HduHgbAR (Table 2, Figure 3). The 3 kb product was gel purified and ligated into pPCR-Script (Table 3, Figure 17A). The resulting 6 kb plasmid was named pHgbA (Table 3, Figure 17B). A signature-tagged transposon Tn1545-Δ3 was PCR amplified from pBluTnSTM1-12 using primers Tn1545K-F and Tn-1545K-R. The 3.8 kb transposon product was gel purified using QIAquick gel extraction kit to remove any impurities. The restriction enzyme *Avr*II was used to linearize pHgbA because it only cut once in the *HgbA* gene to allow the transposon to be cloned into it. Tn1545-Δ3 was ligated into pHgbA and transformed into *E. coli* Top10 cells on LB-Kan (50 μg/ml)- Ap (100μg/ml). The resulting 9 kb plasmid construct was named pHgbA-TnSTM1 (Figure 17C). The STM tag was screened by PCR using primers STM1 and Tn1545R (Figure 17D). For homologous recombination, pHgbA-TnSTM1 was electroporated into *H. ducreyi* (Figure 3). Transformants were screened by colony PCR with STM1 and Tn1545R primers. HgbA transformants were verified by PCR using the HduHgbAF and HduHgbAR primers (Figure 18A). Using those primers, the *hgbA* mutant would produce a larger fragment size of 6 kb due to the insertion of the STM1-Tn1545-Δ3. Three HgbA transformants that were analyzed were Hdu-hgbA-STM1-1, 3 and 8. Figure 18A shows that all the mutants contained the transposon insertion. The mutant genomic DNA was

Figure 17. Agarose gels of hemoglobin receptor A (*hgbA*) cloning into pPCR-Script to construct pHgbA-STM1-Tn1545. (A) PCR amplification of *hgbA* from *H. ducreyi* 35000 genomic DNA using primers HduHgbAF and HduHgbAR. Lanes 1 and 2, show the *hgbA* amplified band. Lane 3 was the 1 kb DNA ladder. (B) Lanes 1 and 2, 6 kb pHgbA plasmid. Lane 3 was the supercoiled ladder. (C) Signature-tagged Tn1545-Δ3 cloned into pHgbA. Lane 1 was plasmid pHgbA-STM1-Tn1545. Lane 2 was the λ *Hind*III ladder. (D) PCR amplification of STM1 tag screening of pHgbA-STM1-Tn1545 using primers STM1 and Tn1545R. Lane 2 was pHgbA-STM1-Tn1545 used as template. Lane 3 was a positive PCR control using pBluTnSTM1 as template and lane 4 was a negative PCR control using dH_2O as template. Lane 1 was a 1 kb DNA ladder. Sizes of the DNA standards were in kilobases.

A



B



C

D

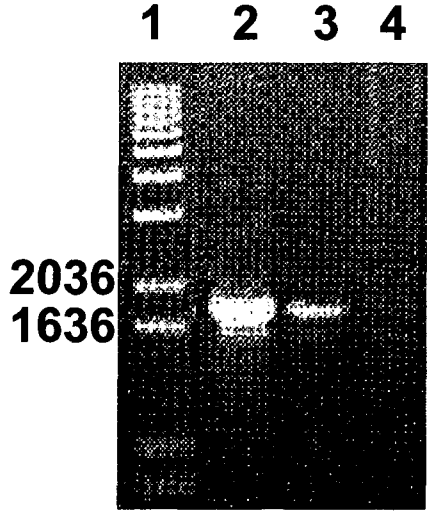
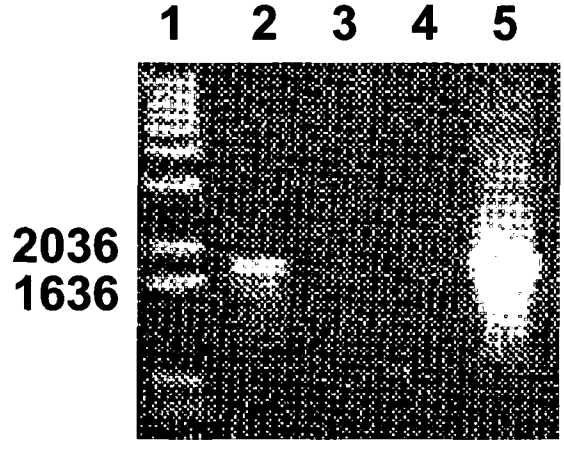


Figure 18. Agarose gels of *H. ducreyi* Tn1545 STM1 tagged *hgbA* mutant screening. (A) PCR amplification to verify insertion of STM1-Tn1545-Δ3 into the *hgbA* gene of *H. ducreyi* 35 000 using primers HduHgbAF and HduHgbAR. Lane 1 was *H. ducreyi* 35000 genomic DNA. Lanes 2, 3 and 4 was genomic DNA extracted from 3 different *hgbA* STM1-Tn1545 *H. ducreyi* mutant clones (Hdu-hgbA-STM1-1, 3, and 8 respectively). Lane 5 was the 1 kb DNA ladder. (B) PCR amplification to verify that the *hgbA* mutant was STM1 tagged using primers STM1 and Tn1545R. Lane 2 was genomic DNA isolated from Hdu-hgbA-STM1-1. Lane 3 was genomic DNA extracted from *H. ducreyi* 35000. Lane 4 was the negative control using dH_2O. Lane 5 was the positive control pBluTnSTM1. Lane 1 was the 1 kb DNA ladder. Sizes of the DNA standards were in kilobases.

A

B



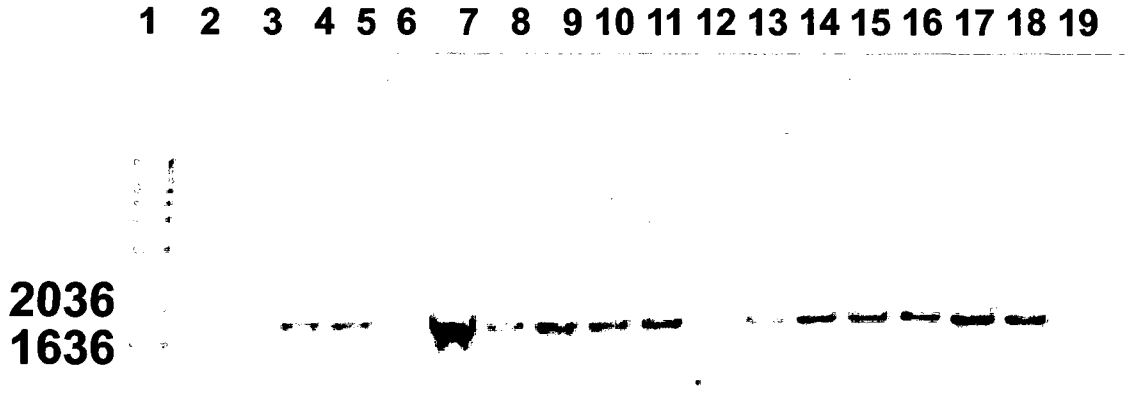
also analyzed by PCR using primers STM1 and Tn1545R to verify that the mutants were STM1 tagged. Figure 18B shows that Hdu-hgbA-STM1-1 was tagged with STM1 resulting in a 2 kb amplicon whereas wild type *H. ducreyi* 35000 was not and did not amplify a band. Sequencing results also confirmed that the *hgbA* mutant was signature-tagged.

3.4 Generation of a Signature-tagged Transposon Mutant Library in *H. ducreyi*:

Electroporation of ST-Tn1545-Δ3 into *H. ducreyi* generated a total of 492 different ST-transposon mutants. Each mutant contained one of the eleven oligonucleotide tags and was colony PCR screened using a STM tag specific primer STM1-12 and Tn1545R that produced a 2 kb fragment. Figure 19 is a representative agarose gel of the STM mutant PCR screening for *H. ducreyi* STM6 and STM7 mutant libraries. Most transformants did contain the STM tag and amplified a 2 kb band (Figure 19). However, they were a few mutants that did not, as shown in Figure 19, lanes 6 and 12. These mutants were discarded and not used in subsequent experiments. *H. ducreyi* STM mutants that had a positive PCR result were stored at -80°C. The mutants were given the nomenclature Hd6-1(20) where the Hd stands for *H. ducreyi*; the first number denotes the STM tag that was incorporated, followed by the mutant number. The number in brackets describes the electroporation date.

Growth curves for a majority of the mutants were performed to assess their growth rate compared to wild type (Figure 17). *H. ducreyi* STM mutants generally followed the typical growth curve trend as wild type, with an initial lag phase, log phase, and stationary phase (OD_{600}). The final OD_{600} for *H. ducreyi* after 24 h never exceeded

Figure 19. Agarose gel of STM tag screening of *H. ducreyi* STM mutants for the generation of transposon mutant libraries. Colony PCR amplification was performed using primers STM6 or STM7 depending on the STM tag incorporated with Tn1545R. Lanes 2-6 were H6-23, 24, 25, 26, and 27 mutants that were screened with primer STM6. Mutant H6-27 did not amplify a 2 kb band (Lane 6). Lane 7 was a positive control using pBluTnSTM6 as template. Lanes 9-18 were H7-1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 mutants that were screened with primer STM7. Mutant H7-4 did not amplify a 2 kb band (Lane 12). Lane 19 was a negative dH_2O control. Lane 1 was a 1 kb DNA ladder. Sizes were in kilobases shown on the left.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
2036
1636

0.5 which, was typical of an *H. ducreyi* growth curve. Mutants that grew noticeably slower than wild type were not included in the STM pools used to infect the TDRM as to not bias the outcome of candidate mutants.

Competition assays were done to assess the attenuation of the mutants *in vitro*. The STM screen identifies mutants that were unable to compete with other strains after mixed infections. The attenuation of selected mutants was confirmed and quantified by competition assays where wild type and mutants were incubated in mixed cultures to rule out comparable growth defects under optimal laboratory growth conditions. For the growth curve OD_{600}, an aliquot from the mixed culture was sampled every 2h for 12h and then at 24h. Figure 20 shows a growth assay that is characteristic of *H. ducreyi* and that the final OD_{600} at 24h does not exceed 0.5. Viable counts were also plated on selective CA-charcoal-Kan (30 μg/ml) and non-selective CA media. The competitive index (CI) was calculated relative to wild type strain 35000. The CI is defined as the output ratio (mutant/wild type) divided by the input ratio (mutant/wild type). A CI = 1 indicated that the mutant grew as well as wild type and does not show any attenuation. A CI less than (<) 1 indicated that the mutant was attenuated in growth compared to wild type (Table 5). The majority of STM mutants showed some attenuation compared to wild type and had CI less than 1 (Table 5). Very few had comparable growth at CI = 1 or CI > 1. Weak attenuation was designated to mutants with a CI in the range of 0.6-1 log difference. Mild attenuation was assigned to mutants with a CI equal to 1-2 log difference and severe attenuation was given to mutants with a CI difference of more than 2 logs. Most of the STM mutants had subtle attenuated phenotypes (CI = 0.1-1 log difference) and did not have significant growth defects *in vitro*.

Figure 20. Competition analysis of *H. ducreyi* STM mutant growth assay of mixed bacterial cultures with wild type *H. ducreyi* 35000 grown for 24 h at 35°C with 5% CO_2. Aliquots were taken every 2 h for 12 h and the OD_{600} was measured. The STM mutants that were examined were H1-2(06), H2-2(18), H7-2(21), H7-8(06), H8-24(16), H10-2(20), H12-3(12) and H12-10(27). Each value represented the mean of three triplicate cultures. Standard deviations of the mean were shown by error bars. Absence of error bars indicated that the standard deviation of the three replicate cultures was too small to be represented on the graph.

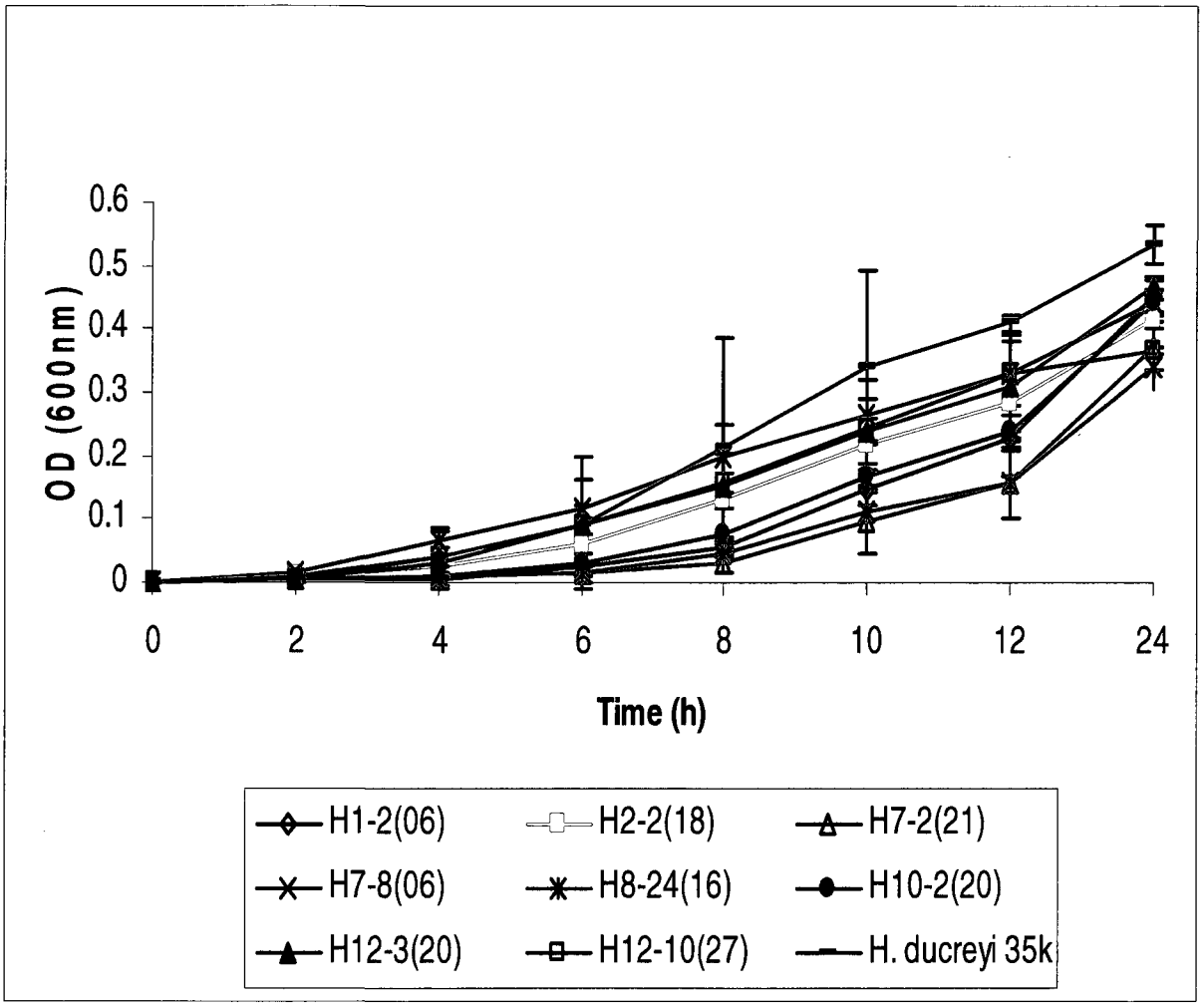

0.6
0.5
0.4
0.3
0.2
0.1
0
OD (600nm)
0
2
4
6
8
10
12
24
Time (h)
H1-2(06)
H2-2(18)
H7-2(21)
H7-8(06)
H8-24(16)
H10-2(20)
H12-3(20)
H12-10(27)
H. ducreyi 35k

Table 5. Competitive indices (CI) were calculated relative to the wild type strain *H. ducreyi* 35000. The competitive index was defined as the output ratio (mutant/wild type) divided by the input ratio (mutant/wild type) at 8h. The data were the average of three independent cultures.

STM mutant strain	Mutated Gene	*In vitro* CI
Hd1-1(24)	*mutS*	0.892
Hd1-2(22)	HD0464	0.644
Hd1-1(22)	*proS*	0.448
Hd1-12(26)	*lpxB*	0.814
Hd1-2(06)	*putP*	0.595
Hd2-2(18)	*ccmH*	0.450
Hd3-5(26)	*recR*	0.847
Hd3-2(26)	*ptrA*	0.188
Hd3-8(26)	HD1268	0.859
Hd4-18(17)	*dnaK*	0.105
Hd5-23(16)	*radA*	0.564
Hd6-8(26)	*cdtA*	0.702
Hd6-31(16)	*norM*	0.562
Hd6-34(16)	*hlyX*	0.423
Hd6-12(16)	*frdA*	0.652
Hd6-30(16)	*proS*	0.327
Hd7-2(21)	*tdhA*	0.887
Hd7-5(13)	*greA*	0.741
Hd7-8(06)	HD1268	0.899
Hd7-2(01)	*sucA*	0.886
Hd8-6(16)	*hgbA*	0.292
Hd8-24(16)	*hlyX*	0.561
Hd9-13(20)	*uspA*	0.485
Hd9-84(24)	*lgtA*	0.357
Hd9-30(20)	*lolA*	0.222
Hd10-2(20)	*hicB*	0.960
Hd12-19(17)	*sodA*	0.708
Hd12-3(20)	*lspA1*	0.836
Hd12-10(27)	*aspS*	0.587

3.5 STM in the Temperature-dependent Rabbit Model (TDRM) of Chancroid (TDRM-STM):

3.5.1 Optimization of the TDRM for STM:

3.5.1.1 Natural History in TDRM:

The natural history of TDRM for chancroid infection was assessed using a pool of STM mutants to confirm its validity. Normally, only one strain of *H. ducreyi* was injected at one time, but in this case a pool of 11 different ST *H. ducreyi* mutants was injected at once. The pool of mutants injected did not show any difference in disease progression compared to wild type. Samples of the inoculum were plated on appropriate agar plates to test for contamination and viability for each experiment. Colony counts of the inoculum were performed to assess the actual amount of bacteria injected which was found to be accurate within the log of CFUs inoculated. Gram stains of the initial inocula also confirmed the presence of *H. ducreyi.*

A sample of heat-killed wild type *H. ducreyi* 35000 HP was also injected on the rabbit's back to assess the inflammatory response. After day 4 of inoculation, the heat-killed sample did not induce inflammation or ulceration (Figure 21A and B). A sample of the media used (GC broth) for the dilutions was also tested and did not result in any inflammation or ulceration by day 4 (Figure 21A and B).

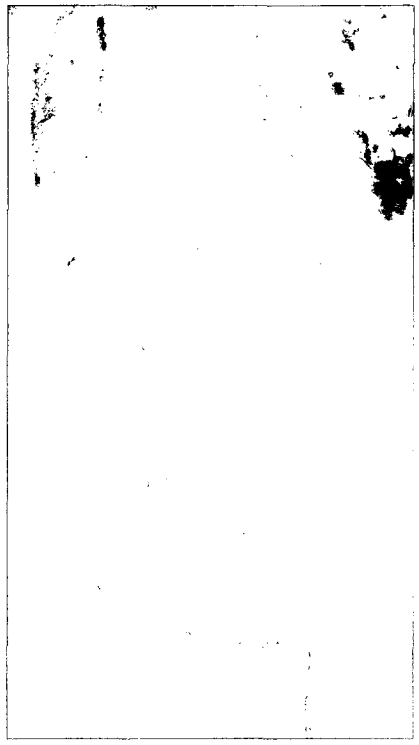
3.5.1.2 Optimization of Inoculum Dose:

The optimal inoculum dose was assessed in the TDRM. For each STM pool tested, a series of dilutions at $10^7 - 10^3$ CFU were tested in duplicate on the backs of the shaved rabbits (Figure 21A). In the same rabbit, wild type *H. ducreyi* 35000 was tested

Figure 21. TDRM-STM1 was performed to determine the natural history of infection and the optimal inoculum dose of STM mutant pools. (A) Appearance of disease 4 days after injection with titered 100 µl inocula containing from 10^7 to 10^3 CFU of 11 uniquely tagged *H. ducreyi* STM mutants in a STM pool in rabbits. Wild type *H. ducreyi* 35000 was injected at 10^4 to 10^3 CFU. HK represents the *H. ducreyi* culture that was heated to 100°C and GC represents the GC media used to prepare dilutions of the bacteria. On the left, is a schematic of the dosages and the sites of injection on the shaved rabbit's back. (B) Appearance of disease 4 days after injection with titered 100 µl inocula containing from 10^8 to 10^4 CFU of *H. ducreyi* 35000 in a control rabbit. W2 was *H. ducreyi* 35000 from a different stock. HK represents the *H. ducreyi* culture that was heated to 100°C and GC represents the GC media used to prepare dilutions of the bacteria. On the left, depicts a schematic of the dosages and the sites of injection on the shaved rabbit's back.

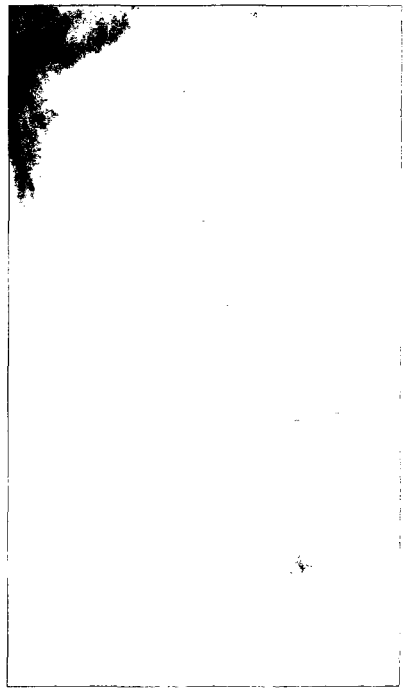
A

STM1 10^7	STM1 10^7	HK
STM1 10^6	STM1 10^6	GC
STM1 10^5	STM1 10^5	W1 10^4
STM1 10^4	STM1 10^4	W1 10^3
STM1 10^3	STM1 10^3	W1 10^4



B

W1 10^8	W1 10^8	HK
W1 10^7	W1 10^7	GC
W1 10^6	W1 10^6	W2 10^4
W1 10^5	W1 10^5	W2 10^8
W1 10^4	W1 10^4	W2 10^4



at 10^4 CFU in duplicate and at 10^3 CFU. In this rabbit, Figure 21A, it was observed that most of the inoculations resulted in lesions with the exception of the injections that were located near the hind legs. However, a second rabbit that was also injected with the same inoculum doses of STM pool 1 and wild type did result in lesions for the same areas (DNS). A third rabbit was injected with a series of dilutions of wild type *H. ducreyi* 35000 at $10^8 - 10^4$ CFU in duplicate. A second sample of wild type *H. ducreyi* 35000 from a different stock was injected at 10^4 CFU in duplicate and at 10^8 CFU. In Figure 21B, the samples that were injected in the second column (middle of the back) demonstrated the most severe response. The second wild type sample injected at 10^8 CFU did not result in severe ulceration but the inoculations at 10^4 CFU were comparable to those observed in wild type 1.

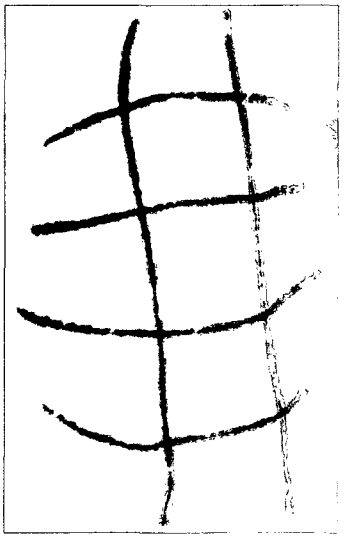
The minimum dose of 10^4 CFU was enough to consistently produce an ulcer comparable to wild type *H. ducreyi* 35000 at the same inoculum. Lesions produced from the inoculums at $10^6 - 10^7$ CFUs were too severe to be consistently used for the animal model. A mixed pool of STM mutants (input pool) was tested at each dilution and demonstrated the same natural history of infection as wild type.

3.5.1.3 Tempo of Ulcer Progression:

The disease progression was consistent with the natural history in the TDRM and in the human challenge model. There were 4 stages of disease progression observed in the TDRM, 1= redness, 2= induration, 3= suppuration, 4= ulceration. Figure 22A is a picture of the rabbit back shaved the night before injection (day -1) with a 3 x 5 grid drawn on its back. Figure 22B shows the rabbit back after the intraepithelial injection of

Figure 22. The appearance of lesions from TDRM-STM1 on day -1, 0, 1, 2, 3, and 4 used to assess the tempo of ulcer progression. (A) Day -1 showed a shaved rabbit back with a 3 x 5 grid drawn on it. (B) Day 0 was the day of inoculation. For each spot on the back of the rabbit, 10^4 CFU of *H. ducreyi* 35000 was injected intraepithelially in a 100 µl inocula. (C) The appearance of lesions on day 1 showed redness and some induration at the sites of injection. (D) By day 2, the lesions appeared as red bumps (induration) and some were filled with pus. (E) The appearance of the lesions on day 3 showed that most of lesions were filled with pus (suppuration). (F) By day 4, the lesions had ulcerated and had started to form scabs over the ulcers.

A



B

C

D

E

F

100 µl of bacteria at 10^4CFU. The injection of 100 µl of culture results in a bubble of inocula just underneath the first layer of epidermis (Figure 22B). After the first day of inoculation, redness was observed at the site of injection (Figure 22C). The following day, the site would consist of induration or a raised area on the skin which was usually filled with pus (suppuration) by the third day (Figure 22DE). Ulceration was usually observed by day 4 (Figure 22F).

3.5.1.4 Lesion Excision and Recovery of *H. ducreyi* STM Mutants:

Since ulceration was usually observed by day 4, the entire lesion was excised from the euthanized rabbits on that day. The excision was performed by the University of Ottawa ACVS staff. Rabbits were first administered an anesthetic to sedate them before they were euthanized. Removal of the lesion was performed by lifting the lesion away from the muscle with tweezers and excising the lesion from the skin. The lesion was placed in a sterile 10 ml specimen jar containing 2 ml of 1X sterile PBS before processing. The tissue was homogenized and 100µL of each suspension was diluted in 1X PBS. The dilutions were plated on CA, CA-charcoal-kan or CA-charcoal-kan-van agar to recover bacteria and to estimate the amount of recovered bacteria from colony counts. *H. ducreyi* has been demonstrated to be resistant to vancomycin and was therefore used to inhibit the growth other organisms found in the skin flora. The increased number of bacteria compared to the input amount from the colony counts demonstrated that the mutants did replicate in the animal model and were thus viable during the experiment (Table 6). Gram stains of the recovered bacteria also confirmed the presence of *H. ducreyi.*

Table 6. Colony counts (as CFU/mL) of bacteria from STM mutant pools in the inoculum and of bacteria recovered from the final ulcer excision. The CFU/mL are the averages of n=2 for 3 different dilutions.

STM POOL #	Initial amount of bacteria from inoculum (CFU/mL ± 0.5)	Final amount of bacteria from ulcer excision (CFU/mL ± 0.5)
1	1.1×10^6	1.4×10^8
2	4.1×10^6	2.3×10^7
3	3.1×10^6	1.1×10^8
4	2.3×10^6	6.1×10^7
5	1.2×10^6	2.5×10^8
6	2.7×10^6	3.3×10^7
7	1.3×10^6	2.5×10^7
8	3.8×10^6	1.5×10^8
9	6.8×10^6	9.8×10^6
10	4.2×10^6	2.1×10^8
11	2.7×10^6	3.9×10^8
12	7.6×10^6	6.4×10^7
13	5.4×10^6	4.4×10^7
14	2.9×10^6	1.3×10^9
15	8.3×10^6	7.1×10^7
16	5.5×10^6	2.9×10^8
17	2.3×10^6	8.5×10^6
18	6.7×10^6	4.7×10^7
19	5.2×10^6	3.2×10^8

3.5.1.5 Histopathology:

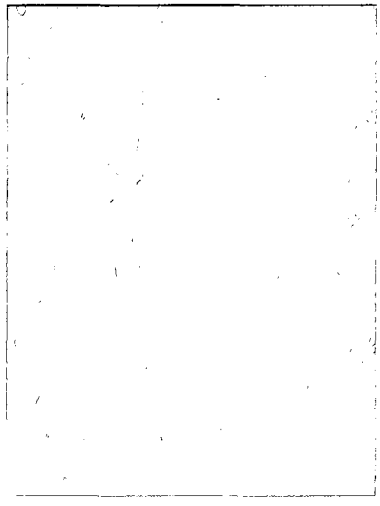
Excised lesions were embedded in paraffin, stained with H&E, and processed by the University of Ottawa Pathology department. The samples were visualized using a Nikon Imaging Microscope and images captured with a digital camera. Consultation with pathologist, Dr. Susan Robertson (Ottawa Hospital) confirmed that the tissue ulceration was a result of bacterial infection. The morphology of the lesions demonstrated the typical tissue damage associated with *H. ducreyi* (Figure 23C). Lesions were acutely inflamed with abundant polymorphonuclear leukocytes and few lymphoid plasma cells with the zonal architecture of chancroid. Three different *H. ducreyi* STM mutant pools were analyzed. The damage to the tissue caused by *H. ducreyi* STM mutant infection was characterized by the extravasation of erythrocytes, extensive edema and tissue destruction that was comparable to wild type (Figure 23D,E and F). The tissue from the control injections of GC media or heat killed *H. ducreyi* was normal without alteration in the epidermal or superficial dermis layer (Figure 23A and B).

3.5.2 *In vivo* Screening of *H. ducreyi* STM Mutants:

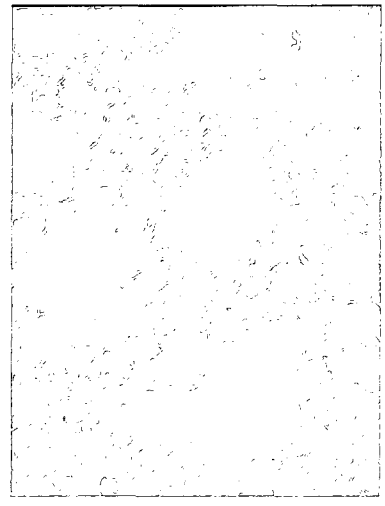
Bacteria recovered from the excised lesions were stored at -80°C and their genomic DNA was extracted for STM tag screening using a Qiagen genomic tip extraction kit. The extracted genomic DNA of the mutants was quantified spectrophotometrically using a Genequant and run on an agarose gel to confirm purity of the sample. The genomic DNA from the STM mutants recovered from the TDRM were screened for STM tags by PCR using STM1-12 and Tn1545R primers that produced a 2

Figure 23. Lesions at day 4 of disease from TDRM-STM3 were stained with hematoxylin and eosin (original magnification x45). Rabbits were inoculated intraepithelially with 100 µl doses of 10^4 CFU of bacteria. (A) control heat killed *H. ducreyi* (B) control GC media (C) *H. ducreyi* 35000 (D) STM pool 1 (E) STM pool 2 (F) STM pool 3.

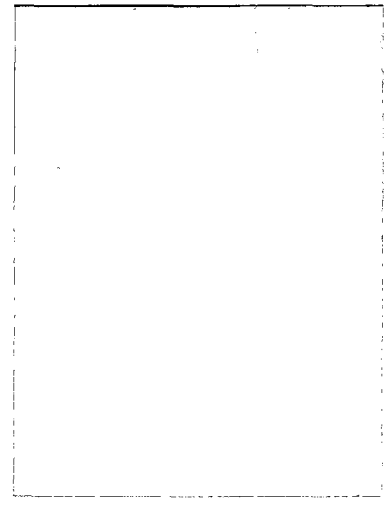
A



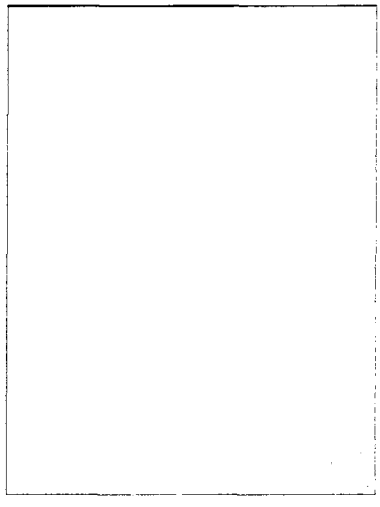
B



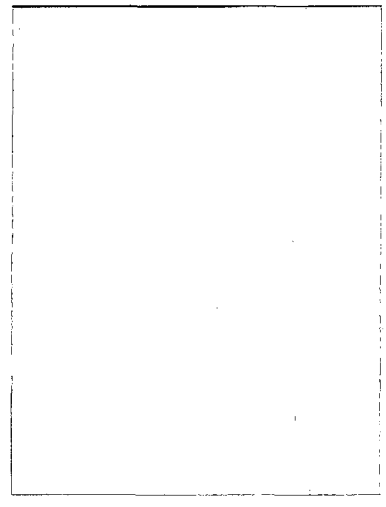
C



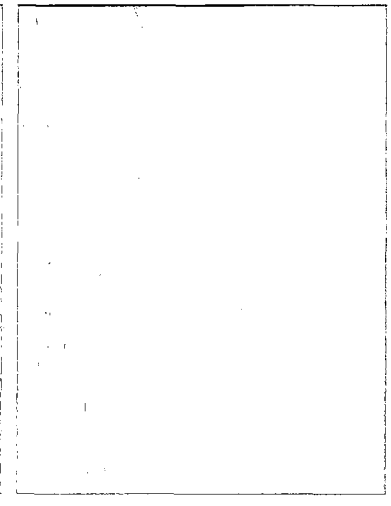
D



E



F



kb band (Figure 24). These mutants were labeled as the output pool (Figure 24B). The genomic DNA consisted of a mixed sample of the STM pool tested and were screened against all the STM primers. Attenuated STM mutant strains were screened by the STM PCR and assessed as strains that did not amplify a 2 kb band with the STM primers in the output pool but were in the initial input pool (Figure 24A). Figure 24 shows the input versus output pool PCR results for STM pool 18 (Table 7). Figure 24B shows the output pool results and that all STM mutants were present with the exception of the STM mutant that was tagged with STM1 and STM8. The mutant with STM tag1 was the STM1-HduHgbA mutant and Hd8-24(16) was the other putative attenuated mutant.

3.5.3 STM1-HduHgbA Positive Control Mutant:

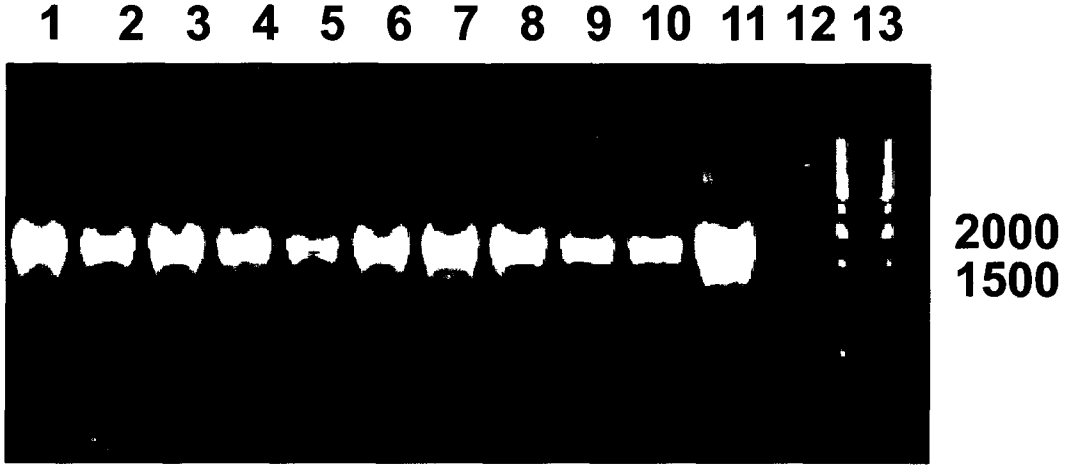
The HgbA mutant has been previously characterized to be avirulent in the TDRM. The inability to recover this ST mutant was used to validate the STM model in *H. ducreyi*. The STM1-HduHgbA mutant was cloned with STM tag 1. It was included in four of the STM pools (STM pool 1, 6, 14 and 18) that were injected in the TDRM and was not recovered in any of the output pools (summarized in Table 7). Figure 24 shows the input and output pool PCR results of STM pool 18 in which the STM1-HduHgbA mutant was included and subsequently not recovered.

3.5.4 TDRM-STM of Chancroid:

Individual STM mutants were selected on the basis of their growth demonstrated in growth assays (Figure 17 and 19) that were comparable to wild type and their *in vitro* competitive index (CI) (Table 5). STM mutants that grew slower *in vitro* compared to

Figure 24. Agarose gels of PCR amplification of STM tags to screen the output pool of recovered bacteria versus the input pool of injected STM mutants. (A) Colony PCR of the input pool of STM pool 18 that was used in TDRM-STM8 using primers STM1-10, 12 and Tn1545R. Lanes 1-10 were mutants STM1-HduHgbA, Hd2-2(22), Hd4-2(13), Hd5-20(16), Hd6-14(22), Hd7-46(20), Hd8-24(16), Hd9-60(20), Hd10-2(26) and Hd12-6(27). Lane 11 was a positive PCR control using pBluTnSTM1 as template and lane 12 was dH_2O. Lane 13 was a 1 kb DNA ladder. (B) PCR amplification of the output pool of STM pool 18 using the same primers described previously. Lanes 1-10 were mutants STM1-HduHgbA, Hd2-2(22), Hd4-2(13), Hd5-20(16), Hd6-14(22), Hd7-46(20), Hd8-24(16), Hd9-60(20), Hd10-2(26) and Hd12-6(27). Lane 11 was a positive PCR control using pBluTnSTM1 as template and lane 12 was dH_2O. Lane 13 was a 1 kb DNA ladder. Sizes of DNA standard were in kilobases.

A


1 2 3 4 5 6 7 8 9 10 11 12 13
2000
1500

B

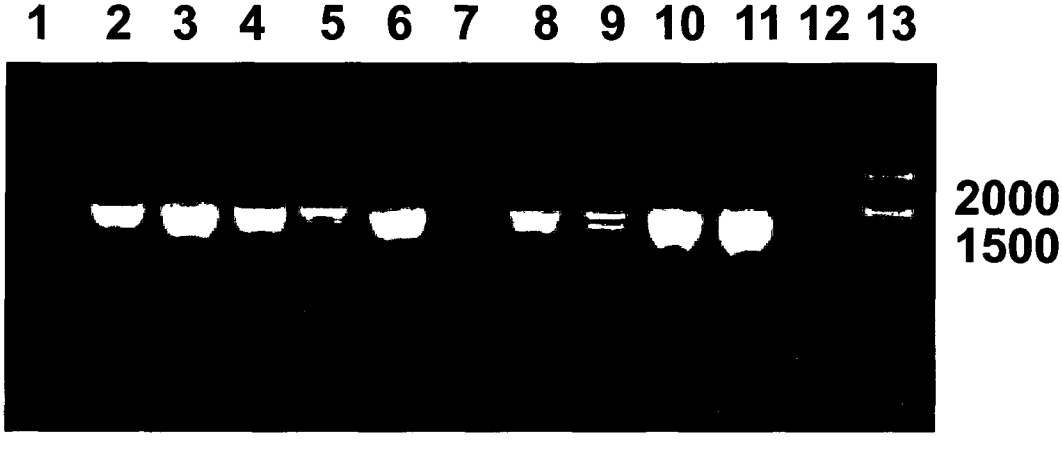

1 2 3 4 5 6 7 8 9 10 11 12 13
2000
1500

Table 7. PCR results of STM tag screening of the input pools versus output pools of the amplified STM tags from *H. ducreyi* STM mutants using primers STM1-12 and Tn1545R. The asterisk (*) denoted when the ST-HduHgbA mutant was included in the experiments. The double asterisks (**) denoted that the mutant Hd8-24(16) was used in two different STM pools.

TDRM-STM #	STM POOL #	INPUT POOL (*H. ducreyi* STM mutant)	OUTPUT POOL (STM#)	PUTATIVE ATTENUATED MUTANTS (STM#)
2	1	STM1-HduHgbA, Hd2-4(06), Hd3-2(06), Hd4-4(16), Hd5-18(16), Hd6-30(16), Hd7-11(13), Hd8-30(13), Hd9-13(20), Hd-10-4(06), Hd12-1(27)	2, 3, 4, 5, 7, 8, 10, 12	**1*, 6, 9**
3	2	Hd1-2(06), Hd2-2 (18), Hd3-2(24), Hd4-1(06), Hd5-17(16), Hd6-12(20), Hd7-2(21), Hd9-66(20), Hd10-19(06), Hd12-15(17)	3, 4, 5, 6, 9, 10, 12	**1, 2, 7**
	3	Hd2-1(06), H3-2(13), Hd4-5(20), Hd5-20(16), Hd6-8(20), Hd7-7(06), Hd8-24(16), Hd9-4(17), Hd10-20(06), Hd12-3(20)	2, 3, 4, 5, 6, 7, 9, 10	**12**
	4	Hd2-3(18), Hd4-3(20), Hd5-22(16), Hd6-16(20), Hd7-12(06), Hd8-34(16), Hd9-2(20), Hd12-1(20)	2, 4, 5, 6, 7, 8, 9, 12	
4	5	Hd1-12(26), Hd2-2(22), Hd3-1(26), Hd4-3(17), Hd5-29(16), Hd6-11(13), Hd7-51(20), Hd8-30(16), Hd9-8(20), Hd10-1(22), Hd12-10(27)	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	**12**
	6	STM1-HduHgbA, Hd4-4(20), Hd6-2(13), Hd7-53(20), Hd8-29(16), Hd9-2(20), Hd12-8(27)	4, 6, 7, 8, 9, 12	**1***
	7	Hd1-2(18), Hd4-36(20), Hd5-2(16), Hd6-4(06), Hd7-57(20), Hd8-2(17), Hd9-12(20), Hd12-7(27)	1, 4, 5, 6, 8, 9, 12	
5	8	Hd1-2(24), Hd3-2(26), Hd4-1(16), Hd5-28(16), Hd6-33(16), Hd7-2(01), Hd8-33(16), Hd9-27(20), Hd10-3(26), Hd12-6(27)	1, 4, 5, 8, 9, 10, 12	**3, 6, 7**
	9	Hd1-1(22), Hd3-4(26), Hd4-3(17), Hd5-12(16), Hd6-33(16), Hd7-4(01), Hd8-6(16), Hd9-29(20), Hd10-7(22), Hd12-5(27)	3, 4, 5, 6, 7, 8, 9, 10, 12	**1**
	10	Hd1-2(22), Hd3-5(26), Hd4-15(27), Hd5-22(16), Hd6-34(16), Hd7-8(06), Hd8-32(16), Hd9-35(20), Hd10-7(26), Hd12-15(17)	4, 5, 9, 10, 12	**1, 3, 6, 7**
6	11	Hd1-5(26), Hd3-5(18), Hd4-3(16), Hd5-18(16), Hd6-1(22), Hd7-5(13), Hd9-84(24), Hd10-13(26), Hd12-5(17)	1, 3, 4, 5, 6, 10, 12	**7, 9**
	12	Hd 4-2(16), Hd5-20(16), Hd6-5(13), Hd7-11(21), Hd8-30(16), Hd9-73(20), Hd10-18(26)	4, 5, 6, 7, 8, 9, 10	
	13	Hd3-7(26), Hd4-1(16), Hd5-23(16), Hd6-8(26), Hd7-45(20), Hd9-3(17), Hd10-2(20)	3, 4, 7, 9	**5, 6, 10**
7	14	STM1-HduHgbA, Hd2-1(22), Hd3-8(26), Hd4-4(27), Hd5-27(16), Hd6-5(16), Hd7-45(20), Hd9-30(20), Hd10-7(22), Hd12-2(27)	2, 4, 5, 6, 7, 10, 12	**1*, 3, 9**
	15	Hd 1-1(24), Hd2-4(26), Hd3-5(18), Hd4-18(17), Hd5-11(16), Hd6-12(16), Hd7-40(20), Hd9-31(20), Hd10-8(22), Hd12-3(26)	2, 3, 5, 7, 9, 10, 12	**1, 4, 6**
	16	Hd1-4(26), Hd2-6(26), Hd3-3(22), Hd4-1(13), Hd6-4(01), Hd7-3(21), Hd8-17(16), Hd9-43(20), Hd10-19(26), Hd12-1(26)	2, 3, 4, 6, 7, 8, 9, 10, 12	
8	17	Hd1-2(22), Hd2-8(26), Hd3-9(26), Hd4-1(27), Hd5-8(16), Hd6-6(22), Hd7-54(20), Hd8-8(16), Hd9-44(20), Hd10-1(26), Hd12-3(27)	1, 2, 3, 4, 5, 6, 7, 9, 10, 12	**8**
	18	STM1-HduHgbA, Hd2-2(22), Hd4-2(13), Hd5-2(16), Hd6-14(22), Hd7-46(20), Hd8-4(16), Hd9-60(20), Hd10-22(26), Hd12-6(27)	2, 4, 5, 6, 7, 9, 10, 12	**1*, 8****
	19	Hd1-12(26), Hd2-4(06), Hd3-1(06), Hd4-4(13), Hd5-3(16), Hd6-16(13), Hd7-48(20), Hd8-17(16), Hd9-61(20), Hd10-10(26), Hd12-19(17)	2, 3, 4, 5, 6, 7, 8, 9, 10	**1, 12**

wild type (at least 2 log difference) were not used in the STM pools. In each TDRM-STM7, three different STM pools were tested at 10^4 CFU and 10^5 CFU in duplicate (Figure 25). Wild type was inoculated at the same dilutions and a negative control of GC broth or a heat killed sample was also injected. Each STM-TDRM trial was tested in two different rabbits in duplicate. Figure 25 shows the appearance of lesions at day 3 from TDRM-STM7 that was injected with STM pools 14, 15 and 16. By day 3, some of the lesions had started to ulcerate. One hundred seventy-six uniquely signature-tagged *H. ducreyi* mutants were tested in 19 different STM pools in a total of 18 rabbits (Table 7).

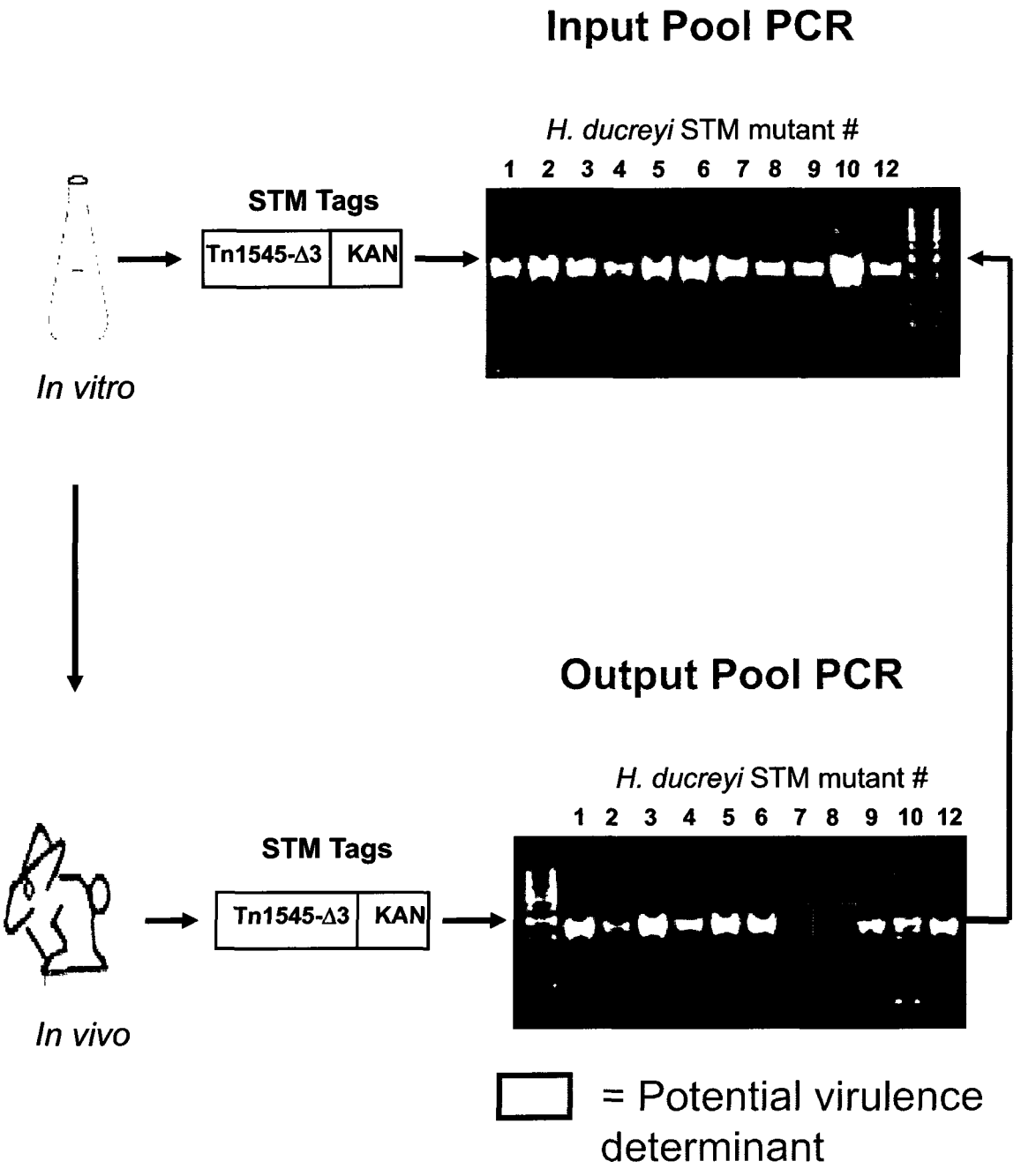
The mixed genomic DNA extracted from the mutants isolated from the lesions (output pool) were screened by PCR using STM1-12 and Tn1545R primers that produced a 2 kb band (Figure 24). The output pool (mutants recovered) was compared to the input pool (mutants inoculated) to assess candidate attenuated STM mutants (Figure 26). Figure 26 represents a schematic of the STM technique to discover potential virulence determinants using PCR amplification of the output pool STM tags and comparing them to the input pool to identify the attenuated mutant. The agarose gels of the PCR results of the output pool represent STM pool 17. The mutant that was tagged with STM8 does not amplify a band. Thus by going back to the input pool, this mutant was identified as Hd8-6(16) and could subsequently be characterized further by sequencing the gene that was knocked out.

Table 7 briefly summarizes the input pool PCR results with the output pool PCR results of all the 19 different STM pools tested. The numbers in the input and output columns represent the STM tag numbers that were present and amplified a 2 kb band. The last column represents the STM tags that were not amplified in the output pool and

Figure 25. Appearance of disease 3 days after injection with titered 100 µl inocula containing from 10^5 to 10^4 CFU of 11 uniquely tagged *H. ducreyi* STM mutants in 3 different STM pools 14 (STM1), 15 (STM2) and 16 (STM3) in rabbits. Each dilution of inoculum was injected in duplicate. Wild type *H. ducreyi* 35 000 was injected at 10^5 to 10^4 CFU. GC represents the GC media used to prepare dilutions of the bacteria. On the left, depicts a schematic of the dosages and the sites of injection on the shaved rabbit back.

GC	STM1 10^5	STM3 10^5
STM3 10^5	STM2 10^4	STM2 10^4
STM2 10^5	STM1 10^4	STM1 10^5
STM1 10^4	W 10^4	STM3 10^4
STM3 10^4	W 10^5	STM2 10^5

Figure 26. A schematic of the *in vivo* STM technique to discover potential virulence determinants using PCR amplification of the output pool STM tags and comparing them to the input pool which was used to identify the attenuated mutant. The agarose gels were of the PCR results of the input and output pool from STM pool 17. The STM tag numbers represented mutants Hd1-2(22), Hd2-8(26), Hd3-9(26), Hd4-1(27), Hd5-18(26), Hd6-6(22), Hd7-54(20), Hd8-6(16), Hd9-44(20), Hd10-1(26), and Hd12-3(27).


Input Pool PCR
H. ducreyi STM mutant #
1 2 3 4 5 6 7 8 9 10 12
STM Tags
Tn1545-Δ3
KAN
In vitro
Output Pool PCR
H. ducreyi STM mutant #
1 2 3 4 5 6 7 8 9 10 12
STM Tags
Tn1545-Δ3
KAN
In vivo
= Potential virulence determinant

thus, signify attenuated mutants to analyze further. The number 1* represents the STM1-HduHgbA mutant that was included in 4 of the STM pools that was injected onto the backs of the rabbits. Each time the STM1-HduHgbA mutant was included, it was not recovered in the output pool (Figure 24B, Table 7). The number 8** represents an *H. ducreyi* STM mutant (Hd8-24(16)) that was included in two different STM pools (STM pool 3 and 18) and was not recovered in either output pool. There were four STM pools tested (STM pool 4, 7, 12 and 16), in which there were no candidate attenuated STM mutants recovered. STM mutants that did not amplify a band from all excised lesions from both rabbits were considered attenuated and selected as candidates for sequencing. In the last column of Table 7, 30 potential attenuated mutants were identified. Twenty-nine candidate mutants were selected for individual genomic DNA extraction and sequencing because Hd8-24(16) was identified twice (Table 7). For the 29 candidate STM mutants, Table 5 also shows their *in vitro* growth defects as calculated as the competitive index (CI) from the competition assays. None of these mutants had major growth defects *in vitro* and had a CI within 1 log difference to wild type.

3.6 The Identification of Virulence Determinants:

3.6.1 Sequencing of Candidate Attenuated *H. ducreyi* STM Mutants:

The candidates were identified from the input pools and their genomic DNA was isolated using a Qiagen Genomic DNA extraction kit as per manufacturer's instruction. The genomic DNA was run on an agarose gel and quantified using a Genequant. Nucleotide sequencing was performed by the University of Ottawa Core Sequencing and at the OHRI-OGIC. Cloned flanking DNA was sequenced using Big Dye (Applied

Biosystems, Foster City, CA) terminators according to the manufacturer's directions and sequence data was generated on an ABI 377 sequencer (ABI Biosystems, Columbia, MD). Sequence alignments and analyses were performed using MacVector (Oxford Molecular, Palo Alto, CA). Similarity searches of DNA were performed by BlastN (GenBank) and for protein sequences using BlastX (SWISS-PROT TREMBL). Searches of the *H. ducreyi* genome Project Database were performed by BlastN. The OHRI-OGIC performed the DNA sequencing and required that the genomic DNA concentration be 50 ng/μl. For samples where this concentration was not achieved, a nested PCR reaction was performed using primers for the transposon and the STM tag. DNA sequencing of the region flanking the transposon was determined using primers for the transposon in an inverse PCR reaction. Sequencing from genomic DNA was difficult and did not work for all the candidates the initial round. Some sequencing reactions worked better than others giving at least 300 bp of sequencing but the majority gave at least 100 bp.

3.6.2 Classification of Mutants:

From the sequencing, the results were organized and classified using Bioinformatics. BlastX searches of GenBank allowed the identification of the genes disrupted by the transposon which was used to identify and compare sequence homology. Table 8 shows the results of the sequencing and the assignment of the genes into eight classes (first column). The second column portrays the *H. ducreyi* STM mutants that were sequenced. The third column depicts the gene identification number (GI) from the sequenced *H. ducreyi* genome in Genbank. The fourth column defines the gene name

Table 8. Candidate *H. ducreyi* virulence genes that were identified by STM. The results were organized and classified using Bioinformatics. BlastX searches of GenBank allowed the identification of the genes disrupted by the transposon which was used to identify and compare sequence homology. Similarity searches of DNA were performed by BlastN (GenBank) and for protein sequences BlastX (SWISS-PROT TREMBL) was used. Searches of the *H. ducreyi* genome Project Database were performed by BlastN. Abbreviation used for bacterial species were: H.i., *Haemophilus influenzae*, E.c., *Escherichia coli*, A.p., *Actinobacillus pleuropneumoniae*, V.p., *Vibrio parahaemolyticus*.

CLASS	STRAIN	GENE ID	GENE	SIMILAR TO (% IDENTITY/ %HOMOLOGY)	PUTATIVE FUNCTION
Classical virulence determinants	Hd6-8(26)	HD0902	*cdtA*		Cytolethal distending toxin subunit A
	Hd10-2(20)	HD0910	*hicB*	H.i.(46/60)	Putative toxin-antitoxin
Cell surface	Hd8-8(16)	HD2025	*hgbA*		Hemoglobin binding receptor
	Hd7-2(21)	HD0388	*tdhA*		TonB-dependent heme receptor
	Hd12-3(20)	HD1156	*lspA1*		Large supernatant protein
	Hd9-84(24)	HD0466	*lgtA*		Cell wall biogenesis
	Hd1-12(26)		*lpxB*	H.i.(74/86)	Lipid A dissacharide synthase
Stress	Hd4-18(17)	HD0189	*dnaK*		Heat shock protein
	Hd12-19(17)	HD0321	*sodA*		Superoxide dismutase A
	Hd9-13(20)	HD1428	*uspA*	A.p.(90/96)	Universal stress protein A
	Hd3-2(26)	HD0187	*ptrA*	E.c.(38/59)	Protease III
DNA recombination and repair	Hd1-1(24)	HD2024	*mutS*	H.i.(71/82)	DNA mismatch repair protein
	Hd5-23(16)	HD1121	*radA*	E.c.(76/89)	DNA repair protein
	Hd3-5(26)	HD0325	*recR*	H.i.(80/90)	Recombinational DNA repair protein
Transport	Hd9-30(20)	HD1237	*tolA*	H.i.(59/71)	Outermembrane lipoprotein carrier
	Hd1-2(06)	HD0999	*putP*	H.i.(67/81)	Sodium/proline symporter
	Hd6-31(16)	HD1428	*norM*	V.p.(43/63)	Multidrug resistant protein
Metabolism	Hd12-10(27)	HD0599	*aspS*	E.c.(72/83)	Aspartyl tRNA synthetase
	Hd1-1(22)	HD1919	*proS*	H.i.(84/91)	Prolyl-tRNA synthetase
	Hd2-2(18)	HD0060	*ccmH*	A.p.(42/59)	Cytochrome c-type biogenesis protein
	Hd6-12(16)	HD0030	*frdA*	E.c.(73/83)	Fumarate reductase flavoprotein subunit A
	Hd7-2(01)	HD1336	*sucA*	H.i.(74/86)	TCA anaerobic respiration
Regulatory	Hd6-34(16)	HD1427	*hlyX*	H.i.(76/90)	Fnr-like transcriptional regulator
	Hd7-5(13)	HD1699	*greA*	H.i.(85/92)	Transcriptional elongation factor
Hypothetical	Hd1-2(22)	HD0464	hypothetical		Putative aminotransferase
	Hd7-8(06)	HD1268	Conserved hypothetical		Conserved putative esterase/lipase

and the fifth column depicts where the insertion mapped in a protein similar to other bacteria, the per cent identity and homology were based on the translated sequence in brackets. The last column describes the function of the expressed gene. The candidates were grouped according to classical virulence factors, putative genes associated with the cell surface, genes in response to stress, genes involved in DNA recombination and repair, genes implicated in transport, genes required for metabolic functions, regulatory genes and genes with hypothetical function.

Within the first class, insertions of the transposon were found in genes that corresponded to classical virulence genes. One has been previously identified in *H. ducreyi*. It was discovered that there were mutations in the *cdtA* cytolethal distending toxin gene (Hd6-8(26)) and *hicB*, a putative toxin-antitoxin cassette (Hd10-2(20)).

The second group included genes that show similarity to those concerned with the cell surface. Insertions were discovered in the *hgbA* hemoglobin receptor gene (Hd8-6(16)), *tdhA* ton-dependent heme receptor gene (Hd7-2(21)), and in the gene coding for the large supernatant protein, *lspA1* (Hd12-3(20)). Hd9-84(24) maps in an *H. ducreyi* gene whose product is involved in cell wall biogenesis (LgtA) and Hd1-12(26) maps to *lpxB* which is involved in anchoring LPS to the outermembrane.

The third class comprises insertions that show similarity to genes involved in response to stress. An insertion was found in the *dnaK* heatshock protein gene (Hd4-18(17)), and the *sodA* superoxide dismutase gene (Hd12-19(17)). Insertions were also found in genes that showed similarity to *uspA*, universal stress protein A (Hd9-13(20)) and in a protease III(*ptrA)* from mutant Hd3-2(26).

The fourth group includes genes that show similarity to those involved in DNA recombination and repair. Specifically genes that were involved in mutation avoidance such as *recR* (Hd3-5(26)), *radA* (Hd5-23(16)), and *mutS* (Hd1-1(24)) were identified.

Three genes associated with transport were assembled in a fifth category. Hd9-30(20) had an insertion mapped to a gene with similar function as an outermembrane lipoprotein carrier (*lolA*), the insertion in Hd1-2(06) was found to be in a sodium/praline symporter (*putP*) and mutant Hd6-31(16) had an insertion in *norM*, a multidrug resistance protein.

Genes classified in the sixth category were those implicated in metabolism. Two insertions were mapped to genes similar to *aspS* (Hd12-10(27)) and *proS* (Hd1-1(22)) belonging to tRNA synthetases. Hd6-12(16) insertion was found to be in a gene coding for a protein similar to a fumarate reductase flavoprotein (*frdA*) and Hd7-2(01) insertion was mapped to a gene also involved in anaerobic respiration (*sucA*). Hd2-2(18) insertion was mapped to a pseudogene for cytochrome c-type biogenesis (*ccmH*).

Two regulatory genes were assembled into the seventh class. Hd6-34(16) showed an insertion in *hlyX*, a Fnr-like transcriptional regulator and Hd7-5(13) showed an insertion in a transcription elongation factor, *greA*.

The last class included two genes with hypothetical functions. One insertion was found in a gene encoding for a hypothetical aminotransferase, Hd1-2(22) and another insertion in Hd7-8(06) mutant was discovered within a hypothetical gene that had a predicted esterase/lipase function.

Three of the mutants that were sequenced were discovered to have insertions in duplicate genes. Hd3-8(26) had an insertion in a transcription elongation factor similar to

Hd7-5(13). The Hd6-29(16) mutant had an insertion mapped to a tRNA synthetase in the same region as Hd1-1(22). Hd8-24(16) had an insertion mapped to the same location as Hd6-34(16) in the Fnr-like transcriptional regulator gene.

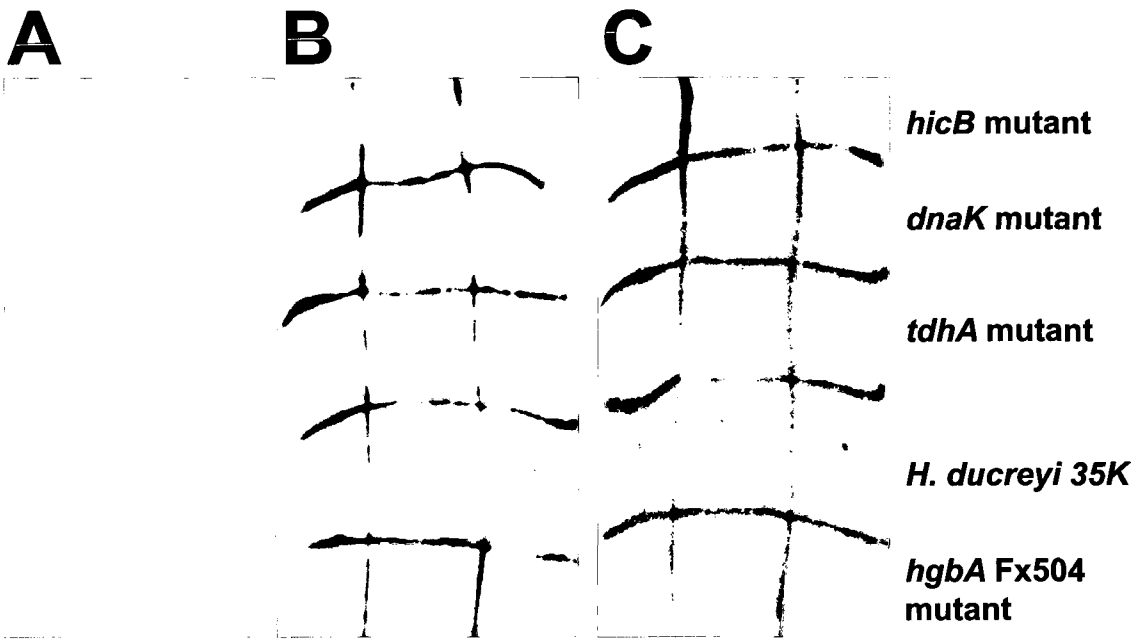
3.7 The Analysis of Potential Mutants:

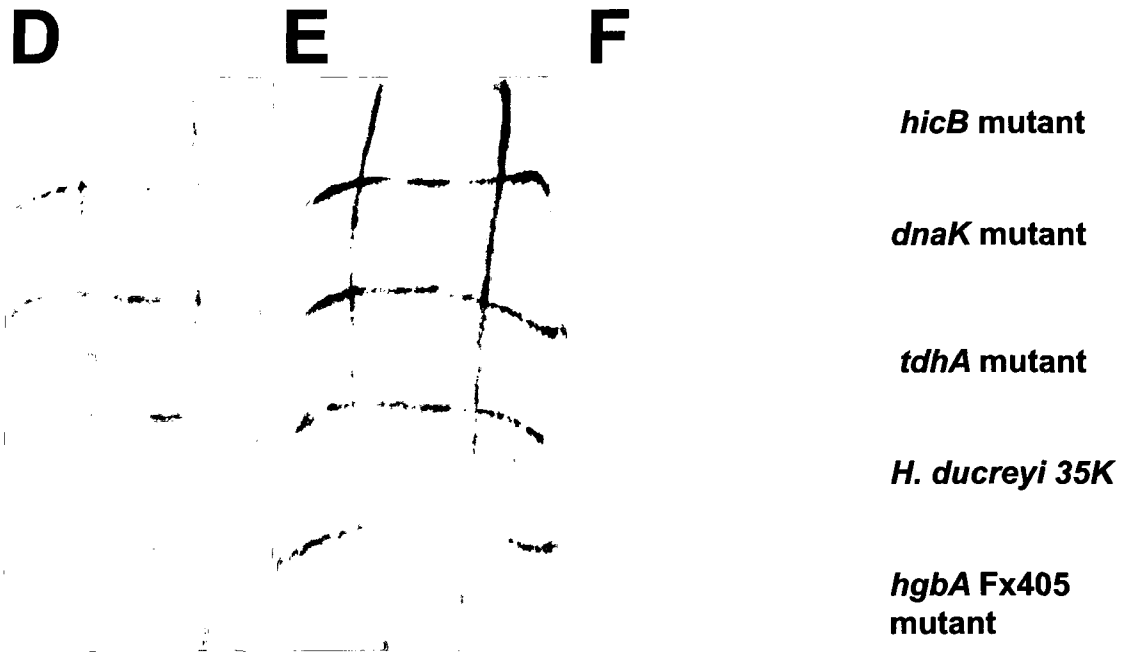
3.7.1 Virulence in the TDRM:

Three candidate *H. ducreyi* STM mutants that were shown to be virulence-attenuated in the STM-TDRM were chosen to be re-screened individually in the animal model to assess virulence. Hd4-18(17) was chosen because it had an insertion mapped to a characterized heat shock protein (*dnaK*) in *H. ducreyi.* Hd7-2(21) had an insertion mapped to a characterized heme receptor protein (*tdhA*) and Hd10-2(20) had an insertion cloned into a gene similar to a putative toxin-antitoxin cassette (*hicB*). All three STM mutants were chosen based on the sequencing results that indicated that these were potentially interesting virulence determinants.

For this experiment, the *H. ducreyi* STM mutants were test for virulence in the rabbits as previously described (38). Each mutant and wild type was inoculated at 10^4 CFU in triplicate on the backs of the rabbit (Figure 27). Three rabbits with the same inoculations were tested for this experiment. An isogenic *H. ducreyi hgbA* (FX504) mutant that was provided by Chris Elkins from the University of North Carolina was used as a control for the rabbit model because it had been previously described as avirulent in the TDRM (135). Two of the injection sites were for measuring the size of the ulcer using microcalipers and observing the ulcer progression by scoring the lesions (Figure 27). The last site was used for culturing of the bacteria to test for the last day of culture

Figure 27. The appearance of lesions after injection with titered 100 µl inocula that contained 10^5 CFU of *H. ducreyi* STM mutants Hd7-2(21) (*tdhA*), Hd4-18(17) (*dnaK*), and Hd10-2(20) (*hicB*) in rabbits. Wild type *H. ducreyi* 35 000 and the *hgbA* Fx504 *H. ducreyi* mutant were injected at 10^5 CFU. Each dilution of inoculum was injected in triplicate. Two of the injection sites were for measuring the size of the ulcer and observing the lesion by assigning a score. The last site was used for culturing of the bacteria to test for the last day of culture positivity. (A) Day 2 (B) Day 4 (C) Day 6 (D) Day 10 (E) Day 12 (F) Day 20


A
B
C
hicB mutant
dnaK mutant
tdhA mutant
H. ducreyi 35K
hgbA Fx504 mutant


D
E
F
hicB mutant
dnaK mutant
tdhA mutant
H. ducreyi 35K
hgbA Fx405 mutant

positivity (Figure 27). The last day of culture positivity was determined after two consecutive tests to plate bacteria from the lesions that did not result in growth. Observations were recorded every two days and the end-point of the experiment was day 20.

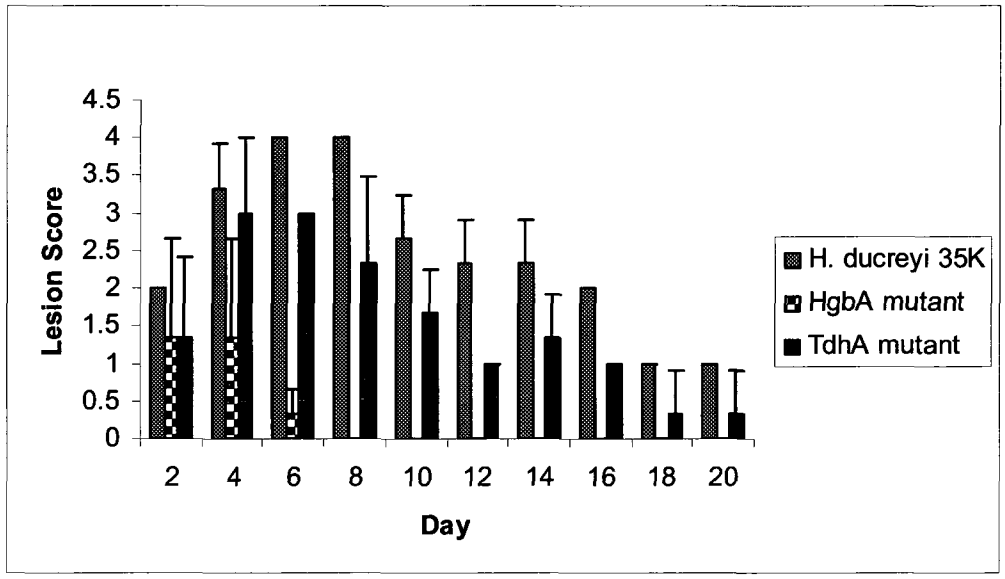
Figure 27 shows the appearance of the lesions for the *tdhA* mutant (Hd7-2(21)), *dnaK* mutant (Hd4-18(17)) and *hicB* mutant (Hd10-2(20)) starting on day 2, day 4, day 6, day 10, day 12 and the end point day 20 results of the individual STM mutant inoculation experiments. Virulence of Hd7-2(21) is summarized in Figure 28 and shows the progression of lesion score in the rabbits. For the data, two lesions were observed and averaged in three rabbits. The *tdhA* Hd7-2(21) mutant was attenuated compared to wild type because the lesions did not reach stage 4 and only developed to the third stage (3±1, $p = 0.1$) which was a pus filled lesion. The Hd7-2(21) mutant lesions were much smaller in size ranging around 5.24 ±1.15 mm ($p = 0.09$) while wild type lesion size averaged around 8.18 ±0.88 mm. The last day of culture positivity for wild type was day 12 ±1 while for the *tdhA* Hd7-2(21) STM mutant it was day 8 ±1. This means that the Hd7-2(21) mutant was unable to survive in the rabbit was as long as wild type. The *dnaK* Hd4-18(17) mutant was also partially attenuated because the lesions did not reach the full ulcerative stage 4 but only reached stage 3 (3.33±0.58, $p = 0.08$) as observed in Figure 29. Compared to wild type, the lesion sizes were much smaller around 5.24 ±1.22 mm ($p = 0.07$). The last day of culture positivity for the *dnaK* Hd4-18(17) STM mutant was day 6 ±1 compared to day 12 ±1 for wild type (Table 9). Finally, the putative *hicB*

Table 9. The last culture-positive day for *H. ducreyi* STM mutants Hd7-2(21) (*tdhA*), Hd4-18(17) (*dnaK*), and Hd10-2(20) (*hicB*), wild type *H. ducreyi* 35 000 and the *hgbA* Fx504 *H. ducreyi* mutant aspirated with 100 µl PBS from lesions on rabbits. Each value represented the mean of three triplicate experiments. Standard deviations of the mean were shown by error bars.

STRAIN	Last Culture-positive Day
H. ducreyi 35000	12 (±1)
hgbA mutant	4(±1)
tdhA STM mutant	6(±1)
dnaK STM mutant	6(±1)
hicB STM mutant	8(±1)

Figure 28. Virulence of *H. ducreyi* STM mutant Hd7-2(21) (*tdhA*) compared to wild type *H. ducreyi* 35000 and Fx504 *hgbA H. ducreyi* mutant at 10^5 CFU in rabbits. The data were for (A) lesion score ($p = 0.1$) and (B) lesion size ($p = 0.09$) for an average of two lesions on three rabbits. Each value represented the mean of three triplicate experiments. Standard deviations of the mean were shown by error bars. Absence of error bars indicated that the standard deviation of the three replicate cultures was too small to be represented on the graph.

A


Lesion Score
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
2
4
6
8
10
12
14
16
18
20
Day
H. ducreyi 35K
HgbA mutant
TdhA mutant

B

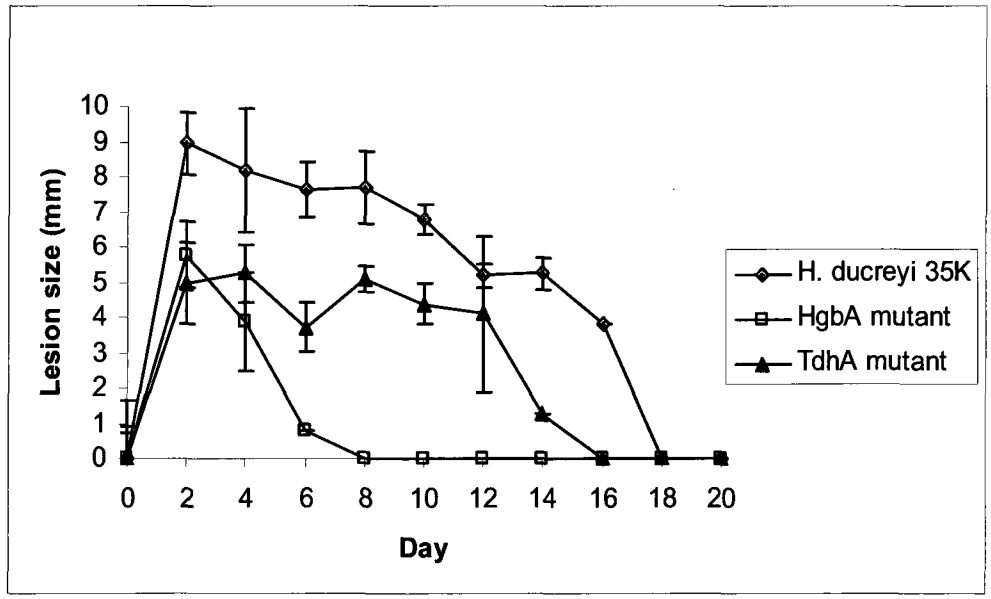
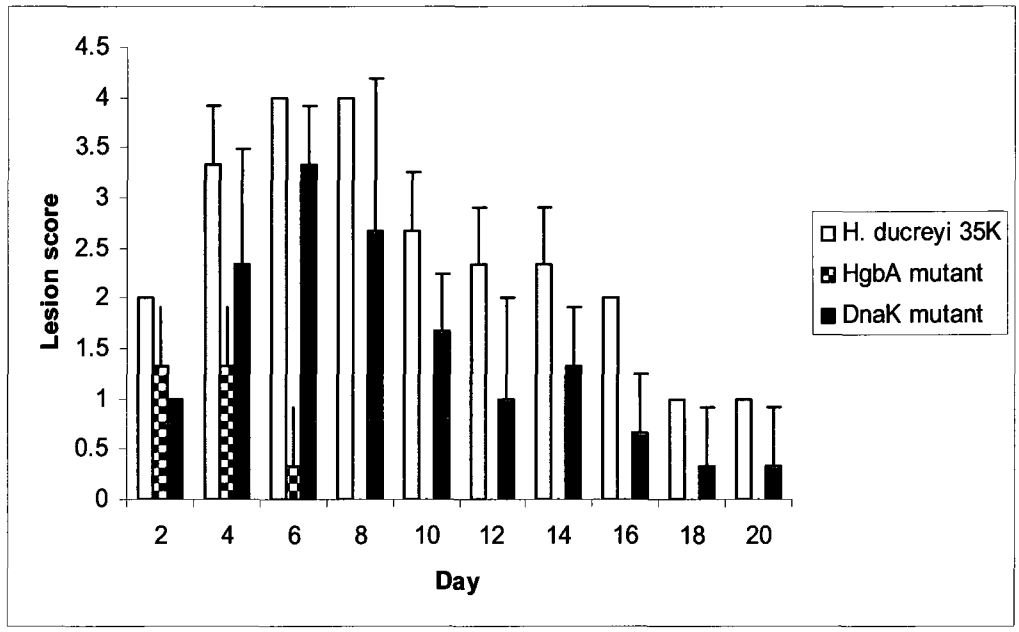
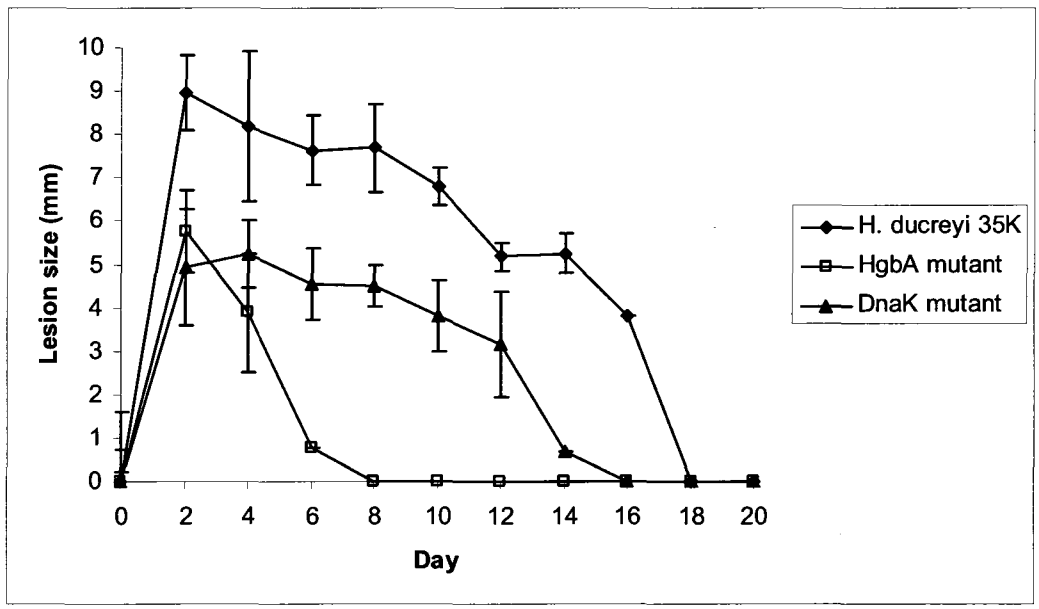

Lesion size (mm)
10
9
8
7
6
5
4
3
2
1
0
0
2
4
6
8
10
12
14
16
18
20
Day
H. ducreyi 35K
HgbA mutant
TdhA mutant

Figure 29. Virulence of *H. ducreyi* STM mutant Hd4-18(17) (*dnaK*) compared to wild type *H. ducreyi* 35000 and Fx504 *hgbA H. ducreyi* mutant at 10^5 CFU in rabbits. The data were for (A) lesion score ($p = 0.08$) and (B) lesion size ($p = 0.07$) for an average of two lesions on three rabbits. Each value represented the mean of three triplicate experiments. Standard deviations of the mean were shown by error bars. Absence of error bars indicated that the standard deviation of the three replicate cultures was too small to be represented on the graph.

A


4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
Lesion score
2
4
6
8
10
12
14
16
18
20
Day
H. ducreyi 35K
HgbA mutant
DnaK mutant

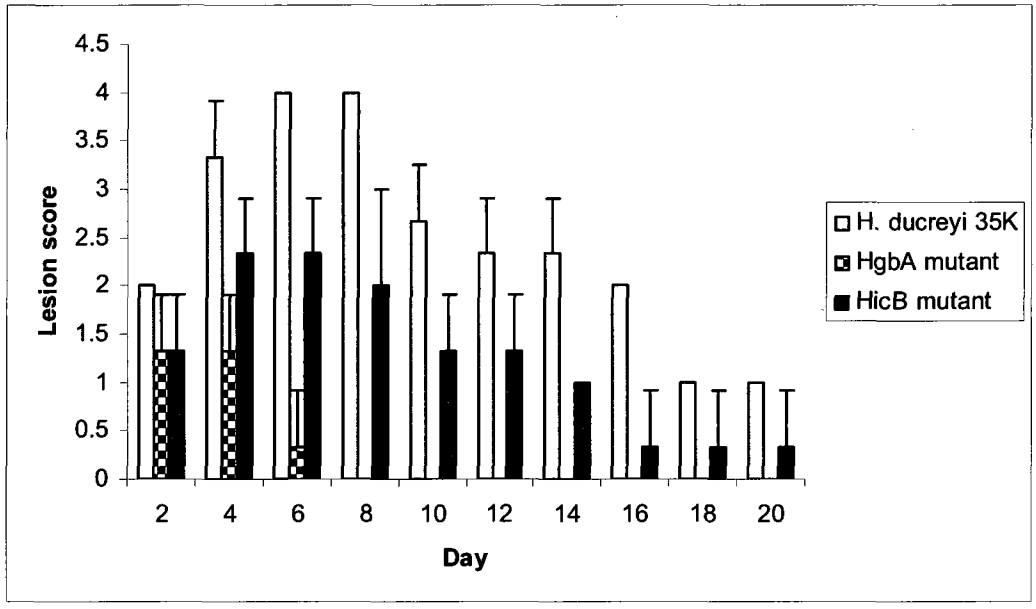
B


10
9
8
7
6
5
4
3
2
1
0
Lesion size (mm)
0
2
4
6
8
10
12
14
16
18
20
Day
H. ducreyi 35K
HgbA mutant
DnaK mutant

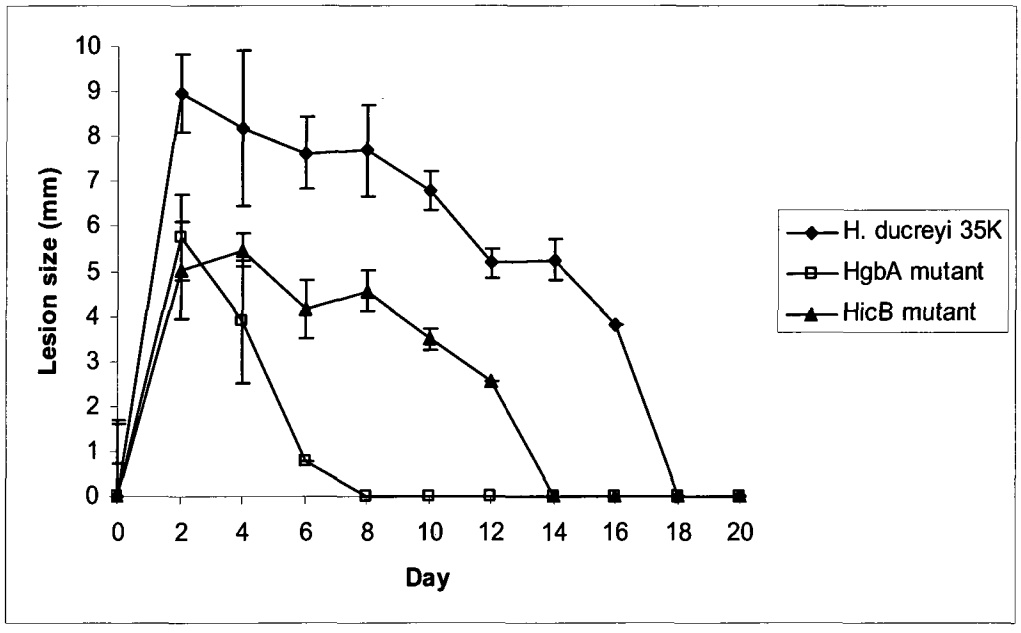
Hd10-2(20) mutant did not grow as well as wild type and was thus partially attenuated (Figure 30). The average size of the lesions was much smaller around 5.48 ±1.09 mm ($p = 0.05$) than wild type (Figure 30B). The lesions only developed to stage 3 (2.83 ±0.58, $p = 0.03$) and did not fully ulcerate (Figure 30A). The last day that the *hicB* Hd10-2(20) STM mutant was culture-positive was on day 8 ±1 (Table 9). The control *hgbA* mutant was avirulent in the rabbit as expected (Figure 30). It demonstrated that it was more attenuated than the candidate STM mutants (Figures 27, 28 and 29). The lesions only developed to stage 2 (1.73 ±0.57, $p < 0.001$) which were characterized by induration and on average was only 3.88 ±0.94 mm ($p < 0.01$) in diameter (Figures 27, 28 and 29). The last day that the mutant was able to be cultured was day 4 ±1 compared to wild type *H. ducreyi* 35000 which was on day 12 ±1 (Table 9).

Figure 30. Virulence of *H. ducreyi* STM mutant Hd10-2(20) (*hicB*) compared to wild type *H. ducreyi* 35000 and Fx504 *hgbA H. ducreyi* mutant at 10^5 CFU in rabbits. The data were for (A) lesion score ($p = 0.03$) and (B) lesion size ($p = 0.05$) for an average of two lesions on three rabbits. Each value represented the mean of three triplicate experiments. Standard deviations of the mean were shown by error bars. Absence of error bars indicated that the standard deviation of the three replicate cultures was too small to be represented on the graph.

A


Lesion score
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
2
4
6
8
10
12
14
16
18
20
Day
H. ducreyi 35K
HgbA mutant
HicB mutant

B


Lesion size (mm)
10
9
8
7
6
5
4
3
2
1
0
0
2
4
6
8
10
12
14
16
18
20
Day
H. ducreyi 35K
HgbA mutant
HicB mutant

CHAPTER 4: DISCUSSION

The signature-tagged mutagenesis (STM) approach is a powerful method that allows large numbers of mutants to be screened for attenuated virulence in an animal model of infection. STM has many advantages over other genomic approaches but the most significant one is that it minimizes the workload and the number of animals required because tagged transposon mutants are pooled and inoculated into the animal host to test in parallel their survival. This study represents the first successful implementation of the STM technique for *H. ducreyi* infection in the temperature-dependent rabbit model. The first part of the discussion describes the proof of principle of STM in *H. ducreyi* while the second part will focus on the putative *H. ducreyi* virulence genes discovered using the STM screen.

Part I Proof of Principle: STM in *H. ducreyi*

The critical components necessary for the transposon-based STM method are: a transposon that has the ability to randomly insert into the pathogen of interest (sections 4.1-2), a signature-tag on the transposon as a way to detect a mutant of interest (section 4.3), and an animal model of infection (sections 4.4-5).

The successful application of STM for a bacterial pathogen requires many important parameters to be addressed within the critical components of STM to demonstrate the validity and reproducibility of the method. During the construction of the signature-tagged transposon, some of the key parameters for successful transposition that were addressed included: evaluation of the STM transposon construct (4.1.1), the random incorporation of the transposon (4.1.2), and the performance of the transposon

(4.1.3). The key parameters for STM pool generation were discussed in the sections on Bacterial fitness (4.2.1), *In vitro* competitive index of the STM mutants (4.2.2), and Summary of STM mutant pools (4.2.3). The key parameters for successful STM tag detection were addressed in section 4.3. The key parameters for a successful *in vivo* model were discussed in the following sections: The natural history of infection (4.4.1), Simultaneous infection (4.4.2), Equal representation of recovered mutants (4.4.3), Advantages of the TDRM (4.4.4), Addressing the TDRM limitations (4.4.5), and The *hgbA* mutant is a positive control for the STM-TDRM (4.4.6). The final *in vivo* proof of principle of STM was addressed in section 4.5: The individual attenuation of virulence for the *dnaK*, *tdhA* and *hicB* mutants. The overall design proof of principle was discussed in the following sections: Saturation of the *H. ducreyi* genome (4.6.1), Complexity of STM pools (4.6.2) and Non-essential genes (4.6.3).

4.1 Key Parameters for Successful Transposition:

4.1.1 Evaluation of the STM Transposon Construct:

A few adaptations of the original STM method were taken during the construction of the signature-tagged transposon vector for *H. ducreyi*. Many critical parameters of the technique were also analyzed to ensure the validity of the STM approach. Transposon Tn1545-Δ3 was chosen as the mutagenic element because it had previously been used successfully to construct hemolytic and LOS mutants in *H. ducreyi* (135, 145). Tn1545-Δ3 was a 3.7 kb transposable element derived from the 25.3 kb transposon Tn1545 that belongs to a family of Tn916 conjugative transposons (34, 94). Contained on the suicide plasmid pMS1 was the transposon Tn1545-Δ3, which was composed of a kanamycin

resistance cassette (*aphA-3* gene encoding 3'-aminoglycoside phosphotransferase type III) and the excision (*xis-Tn*) and integrase (*int-Tn*) genes necessary for the transposition of the element (132). This transposon has also shown its ability to insert randomly and resolutely in the *H. ducreyi* genome (132).

The backbone plasmid pBluescript also served as a suicide vector because it is unable to replicate in *H. ducreyi* and thus for survival purposes, the transposon would have to integrate into the bacterial genome. To verify that the transposon had integrated into the *H. ducreyi* genome and to screen out the plasmid in the *H. ducreyi* transformants, the colonies were counterselected on agar containing the antibiotic resistance selection for the backbone plasmid. For the power of STM to be realized the insertion of the transposon but not whole vector integration into the bacterial genome must be achieved.

4.1.2 Random Incorporation of the Transposon:

Random incorporation of Tn1545-Δ3 into the *H. ducreyi* genome was assessed by Southern blots. The incorporation of the transposon was random based on the probe annealing to different areas on the blot and thus different areas of the genome. A total of 33 mutants were analyzed by Southern blot and three of the mutants had multiple copies of the transposon and two that showed the probe annealing to the same area on the blot. These mutants were not used in the subsequent experiments. All transposons do contain inherent hotspots so it was not unexpected to see these results. However, this problem should not overshadow the results because precautions were taken to ensure that the majority of the mutants were random single insertions of the transposon. To address this problem, more mutants would have to be screened in order to saturate the genome.

4.1.3 Performance of Tn1545-Δ3:

To validate the efficiency of the transposon Tn1545-Δ3, a series of experiments were done, initially, with the pBluTn construct transformants and subsequently in pBluTnSTM1-12 transformants. The stability of the transposon in *H. ducreyi* transformants was assessed and confirmed by serial passage of the mutants. Transposon mutants were still viable at the end of 8 weeks and colony PCR amplification of the kanamycin cassette and the transposon genes confirmed the presence of the transposon in the mutant. It was important to determine if the transposon would be stable and remain in the *H. ducreyi* genome for the duration of the experiments because the TDRM runs for as long as 20 days. In the animal there would not be the selective pressure to retain the kanamycin resistance cassette and thus the transposon for the duration of the experiment.

4.2 Key Parameters for STM Pool Generation:

4.2.1 Bacterial Fitness:

To ensure that mutants recovered from the mutagenesis experiments were not identified on the basis of a slower growth rate than wild type, growth curves of signature-tagged Tn1545-Δ3 *H. ducreyi* mutants were also performed and compared to wild type. *H. ducreyi* STM mutants displayed similar *in vitro* growth characteristics as wild type, with an initial lag phase, log phase, and stationary phase. The growth conditions do not closely reflect *in vivo* growth conditions, but would allow the identification of strains with serious temperature and growth requirements. Most of the strains examined had growth curves that were not significantly different than those of wild type. This was important for creating STM pools that would be used to inoculate the animal model.

Mutants that grew at a much slower rate (at least 1-2 log difference) compared to wild type were not used in the pool of mutants to be tested in the TDRM. A slower growing mutant could potentially be misinterpreted as a false positive for STM. Slower growing mutants could still be virulent in the TDRM but take longer for disease progression.

4.2.2 Competitive Index of the STM mutants:

Competition assays were performed to assign an *in vitro* phenotype corresponding to the attenuation of virulence in the mutants. These competition assays of the mutant versus wild type were also done to evaluate the effect of dominant strain or compensation within a pool of mutants. The STM screen identifies mutants that are unable to compete with other strains after mixed infections. The attenuation of selected mutants was confirmed and quantified by competition assays where wild type and mutants were incubated in mixed cultures to rule out comparable growth defects under optimal laboratory growth conditions. The competitive indices (CI) are calculated relative to wild type strain *H. ducreyi* 35000. A CI = 1 indicated that the mutant grew as well as wild type and did not show any growth defects. The majority of STM mutants showed some attenuation compared to wild type and had CI less than 1 (Table 5). Very few mutants had comparable growth at CI = 1 or CI > 1. The selection of the mutants to be included into a pool was partially based on their *in vitro* CI. Mutants with a CI = 1 or close to 1 were chosen because they represented strains that differed slightly in growth compared to the wild type parent strain *in vitro* but did not have significant overall growth defects. Some mutants with a low CI (CI > 0.2) were included in some of the pools to analyze their growth *in vivo*. Most of the mutants were considered to have weak to mild

attenuation *in vitro*. Weak attenuation was assigned to mutants with a CI in the range of 0.6-1 log difference. Mild attenuation was designated to mutants with a CI equal to 1-2 log difference and severe attenuation was given to mutants with a CI difference of more than 2 logs. Out of the 30 mutants with low CI screened, only two were identified as potential attenuated candidates. These mutants had insertions in the *dnaK* and *ptrA* gene. The rest of the mutants with low CI did not have a growth disadvantage *in vivo* and were pathogenic. Overall, most of the STM mutants had subtle attenuated phenotypes and did not have significant growth defects *in vitro*. This observation was also observed in other STM screens (28). However, *in vivo*, the STM mutants showed a greater degree of attenuation compared to the wild type parent (discussed in subsequent section). The decreased survival of these mutants in the rabbit could be due to their defective growth *in vivo* and/or increased susceptibility to innate immune mechanisms of the host.

4.2.3 Summary of STM Mutant Pools:

A library of 492 unique STM mutants of *H. ducreyi* was constructed by using 11 uniquely tagged Tn1545-Δ3 transposons. No transformants for STM11 were obtained and thus were not used in the subsequent STM pools. A few transformants that did grow on the selection agar did not amplify the STM11 tag. The reasons for this are not clear. It could be due to spontaneous or spurious mutants that grow on kanamycin without the transposon or it could have been due to the PCR amplification process. Before transformation with the STM11 construct, the tag was amplified using the STM primers but perhaps the transformation disrupted the genes necessary for transposition or for tag amplification. Each mutant contained one of the eleven oligonucleotide tags and was

colony PCR screened using a STM tag specific primer and a transposon specific primer. Every STM mutant selected for a pool had previously been screened by growth assay and competition assay. Figure 23 is an overview of the STM technique used to discover potential virulence determinants using PCR amplification of the output pool STM tags and comparing them to the input pool to identify the attenuated mutant.

4.3 Key Parameters for Successful STM Tag Detection:

PCR-based STM has been used to successfully identify and screen *Pseudomonas aeruginosa* (77), *Actinobacillus suis* (98), *Salmonella gallinarum* (122), *Francisella tularensis* (137), *Edwardsiella ictaluri* (141) and *Mesorhizobium loti* (124) mutants. The STM tags used were specifically designed for PCR and were adapted from Lehoux (77), refer to Table 4. The STM oligonucleotides that were synthesized as tags, were each 21 nucleotides long with a T_m of 64°C that allows for one step PCRs when primer combinations are used for screening. The consensus 5'-ends are comprised of the first 13 nucleotides and have higher ΔGs for optimizing PCRs. The variable 3'-ends indicated by the underline, define tag specificity and permit amplification of specific DNA fragments. These tags facilitate single step PCR amplification for the screening of mutant pools, by using each tag as a forward primer in conjunction with a reverse primer synthesized within the Tn1545-Δ3 transposon (Tn1545R). Since Lehoux *et al* had used the PCR STM tags in previous experiments, the suitability of the individual tags for detection within a pool of differently tagged mutants had already been tested. However, the ability to detect the tags in the input and output pools would need to be tested and confirmed for *H. ducreyi.* The 12 unique signature tags (77) were cloned adjacent to the transposon

Tn1545-Δ3 using inverse PCR. Specificity of the ST primer was demonstrated because each construct was tested with each STM primer in a mixed PCR and only the STM primer that corresponded to the ST that was cloned into the vector produced an amplicon. This was important for STM screening of the mutants after injecting into the animal model to avoid cross reactivity of the tags. Consequently, this is the main advantage of PCR-STM because the analysis of the PCR products directly indicates the presence or absence of the tags.

4.3.1 Addressing the PCR-STM Limitations:

The limit of PCR-STM is that it is less quantitative then other tag detection systems such as hybridization. Using hybridization of input tags to output tags gives an indication of the degree of attenuation because weak signals can also be detected. The use of real-time PCR has been used as a detection method for an STM screen in *Burkholderia cenocepacia* (58). Real-time PCR permitted the relative quantification of template DNA from each mutant which allowed the degree of attenuation to be observed. Another limitation is that PCR-STM requires a large number of PCR products that need to be analyzed by gel electrophoresis.

To address this problem, a recent STM screen of *P. aeruginosa*, used a multiplex PCR-based STM to screen 72 mutants with a tag-pool size of 24 PCR based tags (111). This modification used a small number of tags with known sequences that had been combined with three different selection markers. The recovered mutants were identified by PCR amplification, in which a tag-specific primer was combined with three primers

that anneal to the selection marker-specific primer instead of one generic primer. This produced three different PCR products and cut down the amount of gels to be run.

4.4 Key Parameters for a Successful *in vivo* Model:

4.4.1 Natural History of Infection:

Preliminary experiments were performed to standardize the infection dose and duration of infection time, and to test the reproducibility of the STM screen. The natural history of the chancroid animal model was also assessed to ensure that the model was reproducible when used for a different set of experiments. Compared to wild type, the pool of mutants injected did not show any difference in disease presentation or tempo and was consistent with the natural history in the TDRM and in the first two weeks of infection in the human challenge model. There were 4 stages of disease observed in the TDRM. The first stage is characterized by redness, the second stage is induration, the third stage is suppuration, and the fourth stage culminates in ulceration. After the first day of inoculation, redness was observed at the site of injection. The following day, the site would consist of induration or a raised area on the skin which usually filled with pus (suppuration) by the third day and ulceration was usually observed by day 4. By day 6 and 8, some of the lesions were getting smaller in size and starting to resolve. Accordingly, for the STM mutant screen, day 4 was chosen as the endpoint when the rabbits were euthanized and the entire lesion was excised in order to recover the full repertoire of mutants. If the infection model was prolonged there could be a chance of a false positive if the mutant did cause ulceration but died off earlier then wild type. The

PCR amplicon was present in the recovered bacteria and the STM tags were well represented and consistent in all samples at day 4.

Histopathology of the lesions revealed that the tissue damage was a result of bacterial infection. The morphology of the lesions demonstrated the typical tissue damage associated with *H. ducreyi*. Lesions were acutely inflamed with abundant polymorphonuclear leukocytes and few lymphoid plasma cells with the zonal architecture of chancroid. Three different *H. ducreyi* STM mutant pools were analyzed. The damage to the tissue caused by *H. ducreyi* STM mutant infection was characterized by the extravasation of erythrocytes, extensive edema and tissue destruction that was comparable to wild type. The tissue from the control injections of GC media or heat killed *H. ducreyi* was normal without alteration in the epidermal or superficial dermis layer. All the lesion samples were analyzed and confirmed with pathologist, Dr. Susan Robertson (Ottawa Hospital).

4.4.2 Simultaneous Infection:

One of the critical parameters in STM is dependent on the ability to establish an infection simultaneously, meaning that all the mutants in a pool used to infect a single animal must be present in sufficient numbers. This generally means that large doses must be used in infections studies. The optimal inoculum dose of STM mutants was standardized and assessed in the TDRM. The minimum dose of 10^5 CFU for a mixed pool of STM mutants (input pool) was enough to consistently produce an ulcer comparable to wild type 35000 HP at the same inoculum. This inoculum dose is

consistent with the amount used in previous experiments of *H. ducreyi* infection in the TDRM (37).

4.4.3 Equal Representation of Recovered Mutants:

Another key requirement for successful STM screening is the ability to recover sufficient numbers of inoculated bacteria from the infection model to allow the identification of each representative mutant that survives within the mutant pool. Colony counts and PCR screening of the mutants recovered from the lesions (output pool) was consistent in all of the inoculums at different dilutions. The larger colony counts of recovered bacteria compared to the initial inoculum amount demonstrated that the pathogens were able to replicate in the infection model. The PCR amplicon was present in the recovered bacteria and the STM tags were well represented. There did not seem to be any bias towards the recovery of tags. One STM tag was not favoured over another in the output pools and were all PCR amplified.

4.4.4 Advantages of the TDRM for STM:

The advantage of using the TDRM for STM was that multiple mutant pools could be tested in one rabbit. This minimizes the number of animals used and screens more mutants at one time. The nature of the infection model which consists of isolated lesions on the back of the rabbit allows a maximum of fifteen different injection sites. However, some areas on the rabbit back are better for infection then others and variation of disease severity between rabbits injected with the same inoculum occurs. Slight variation in disease presentation is expected when dealing with different rabbits due to differences in

their immune responses. This genetic diversity of the rabbits is also considered another advantage of the TDRM. It was observed that most of the inoculations resulted in lesions with the exception of the injections that were located near the hind legs in one rabbit. A second rabbit that was also injected with the same inoculum doses of an STM pool and wild type did result in lesions for all sites of injection. Thus, each STM pool was tested in two different rabbits and in quadruplicate in each rabbit. Only mutants that were not recovered in any of the rabbits from all the lesions from each site of injection and therefore did not produce a PCR amplicon in the output pool were retained for further analysis. These mutants were putatively assigned attenuation in virulence.

4.4.5 Addressing the TDRM Limitations:

The *in vivo* model itself may alter the efficiency and reproducibility of the screening procedure. Some tags are not amplified from the chromosomal DNA despite all the precautions taken during the initial tag selection. It is an unexplained phenomenon but does occur frequently. This problem of signal quality affects the reproducibility and can lead to the identification of "false-negative" candidates. To ensure that false negative candidates are not selected, two rounds of screening has been performed by many groups doing STM studies (35). In this experiment, pools of mutants are inoculated in an initial screen tested in the animal model. Attenuated candidate strains were identified because they did not generate a PCR amplicon in the output pool compared to the input pool. These attenuated strains were reassembled into a new pool and inoculated into the animal model. In the second round, these assembled mutants were screened in duplicate and the mutants with reproducible absence of PCR amplicons were identified. This was not used

in the STM study for *H. ducreyi*, mainly due to time constraints and further investigations to improve the reproducibility of our screen were not undertaken. However, some experiments were done to test the reproducibility and reliability of the STM screen. STM pool 1 was tested in two different experiments. Thus, these mutants were screened twice and the same mutants that could not amplify a PCR band in the first trial were observed in the second. An individually tagged *H. ducreyi* STM mutant (Hd8-24(16)) was included in two different STM pools (STM pool 3 and 18) and was tested in four different rabbits. This mutant was not recovered in either output pool. Table 7 briefly summarizes the input pool PCR results with the output pool PCR results of all the 19 different STM pools tested. The last column represents the STM tags that were not amplified in the output pool and thus were potential attenuated mutants to analyze further.

4.4.6 The *hgbA* Mutant is a Positive Control for the TDRM-STM:

The STM1-HduHgbA mutant was constructed to prove the reliability of the STM screen because the isogenic hemoglobin A mutant (FX504) has been previously characterized in the rabbit model to be attenuated (135). Therefore, the STM1-HduHgbA mutant would serve as a positive control the STM screening of *H. ducreyi* mutants in the TDRM. The STM1-HduHgbA mutant was included in 4 of the STM pools that were injected in the rabbits. Each time the STM1-HduHgbA mutant was included, it was not recovered in the output pool (Figure 21B, Table 7). Thus the inability to recover this mutant from each specific pool confirmed the sensitivity and specificity of the model and STM screen.

4.5 Final *in vivo* Proof of Principle:

Three potentially attenuated mutants were used in a challenge experiment to verify and determine the degree of attenuation caused by the transposon mutation. Initial screening was performed by injecting a pool of eleven mutants at one site on the back of the rabbit. Within the pool, mutants may not grow as well or their attenuation may be compensated and complemented by other mutants. The putative attenuated mutant strains recovered need to have their virulence phenotypes confirmed individually in the animal model.

4.5.1 Individual Attenuation of Virulence for the *dnaK, tdhA and hicB* Mutants:

Three candidate *H. ducreyi* STM mutants that were shown to be virulence-attenuated in the STM-TDRM were chosen to be re-screened individually in the animal model to assess virulence. Hd4-18(17) was chosen because it had an insertion mapped to a characterized heat shock protein (*dnaK*) in *H. ducreyi.* Hd7-2(21) had an insertion mapped to an uncharacterized heme receptor protein (*tdhA*) and Hd10-2(20) had an insertion cloned into a gene similar to a putative toxin-antitoxin cassette (*hicB*). All three STM mutants were chosen based on the sequencing results that indicated that these were potentially interesting virulence determinants.

For this experiment, the *H. ducreyi* STM mutants were tested for virulence in the rabbits as previously described (37). For virulence analysis, each mutant is compared to the wild type parent based on lesion score, lesion size and the last day of culture positivity. There are three categories that the mutants can be assigned: virulent, defined as those that ulcerate to the same extent and severity as wild type *H. ducreyi* 35000;

attenuated, those that do not form pustules and resolve within 3 days; and partially attenuated, those that form lesions that do not fully ulcerate or are smaller in size and reach the last day of culture positivity in less time then wild type.

All three of the candidates demonstrated partial attenuation in virulence. The *tdhA*, *dnaK* and *hicB* mutants only developed pustules in the rabbits and did not form ulcers to the same degree as wild type. The lesions formed by the mutants were only half the size as those from the wild type parent and the mutants were unable to survive in the rabbit as long as the wild type strain. The control *hgbA* mutant was avirulent in the rabbit as expected and was more attenuated than the candidate STM mutants. The lesions formed by the hgbA mutant only developed into papules and resolved by the fourth day of the experiment. These data suggest that the candidate genes were not required for full virulence but that each were moderately important for survival in the animal model. This confirms the reliability of the STM screen to discover attenuated mutants.

The individual inoculation of the candidate mutant in the animal model to determine the degree of attenuation is a direct method for this assessment. The degree of attenuation of *H. ducreyi* infection is ascertained from observable clinical signs. Other *H. ducreyi* mutants have been tested before in the same manner and this ensures the reproducibility of the model for STM. Another way to determine the degree of attenuated virulence of the putative mutants is to perform *in vivo* competition assays. This method is usually employed in STM studies when the pathogen infects internal organs, such as the liver, spleen and lungs that do not have observable clinical features. Furthermore, reduction of LD_{50} values can be calculated and compared to those of the wild type parent in order to determine the degree of attenuated virulence. Even though

the criteria chosen to define an attenuated candidate can vary between different studies, these findings provide a means to assess the first step in the characterization of novel virulence genes.

4.6 Proof of Principle STM Summary: The Overall Design

4.6.1 Saturation of *H. ducreyi* Genome:

The predicted size of the *H. ducreyi* genome is 1.7Mb which encodes approximately 1800 ORFs. Thus, the 176 STM mutants screened to date only represent about 1% of the genome covered with the potential to cover more of the genome. Previous STM studies have shown approximately 5-10% of the total number of mutants will be negatively selected by STM (108). Not all of the STM mutants were tested in the rabbit model due to the limited time frame for these experiments. These mutants can still be screened in the TDRM to identify interesting virulence determinants and increase the number of mutants that may be negatively selected. To saturate the genome, more STM *H. ducreyi* mutants should be constructed to increase the numbers in the library.

4.6.2 Complexity of STM Pools:

The pool of STM mutants in other studies usually ranges from 48 to 96 mutants per inoculation (117). A small pool size of 9 STM mutants of *E. ictaluri* were injected into catfish and did not have a negative affect on the outcome of the study (141). In total 1071 mutants were analyzed and 50 attenuated candidate strains (about 4.7%) were identified, which falls in the expected range of the number of mutants identified. Another improvement that could be done to this STM study would be to increase the complexity

of the pools. To increase the complexity of the injected pools, more PCR STM tags could have been designed to allow more individually tagged mutants to be screened within a given pool. Kavermann *et. al.* used 24 STM PCR tagged *H. pylori* mutants per pool to infect gerbils. They screened a total of 960 mutants that resulted in the identification of 47 potential virulence genes and discovered a novel secreted collagenase (70). However, there is a restriction as to how many different mutants for use as inocula in infection studies can be used. As the complexity of the pool increases, the probability that some virulent mutants will not be present in sufficient numbers also increases, which would lead to false identification of avirulent mutants. Thus a balance must be achieved when determining pool size and complexity.

4.6.3 Addressing STM Limitations:

STM is limited to finding non-essential genes, specifically genes that are not required for growth in broth. The mutants must be able to survive *in vitro* so that they can be assembled in pools for injection into an animal host. Another short coming is the inability to detect whether a certain gene is dispensable for growth during the negative selection or whether it is simply not present in the mutant library. To assess this problem, the transposon insertion sites for all the *N. meningitidis* mutants were mapped (52). However informative this experiment was, it is time consuming and labour intensive to be performed for every STM analysis. Nevertheless, in the interest in saving time, the output of interesting mutant strains can be used directly for additional investigations. The mutant is readily accessible for further functional characterization. This is an important

point because there are numerous genes that have yet be given a precise role in pathogenesis.

A review of all the genes identified in STM studies reveals that there are few overlaps between different screens, even when STM analysis is performed on the same bacterium a number of different times. This implies that it is difficult to be completely exhaustive. In part, this is due to the inherent variability of the animal model used for screening. The route of infection and the target organ will have an impact on the type of candidates identified. Little overlap and redundancy of candidate genes identified can also be caused by the nature of transposable elements. Transposons have inherent hotspots and do not entirely have random incorporation into the genome. In addition, there is the possibility of the polar effect of the transposon on downstream genes. In the initial screening this could be advantageous to observe genes encoding the same function. This problem can be easily addressed by constructing an in-frame chromosomal deletion of the gene of interest and also by complementation of the attenuated phenotype of the candidate gene. In spite of all the experimental difficulties and limitations, STM has been applied successfully to many pathogens. It is a powerful and robust high-throughput screening technique for the analysis of genes that are not essential for life but are needed for growth in specific environments. Novel genes have been discovered that provide new insights into mechanisms of pathogen persistence within a host. In particular, new understanding on genes involved in host tropism, tissue-specific tropism, adaptation to stress, nutritional deficiencies and extracellular or intracellular life styles have been uncovered.

CHAPTER 5: DISCUSSION

Part II Relevance of Findings: Literature Review of Identified Genes

From the sequencing, the results were organized and classified using Bioinformatics. BlastX searches of GenBank allowed the identification of the genes disrupted by the transposon which was used to identify and compare sequence homology. Table 8 is a summary of sequencing and the assignment of the genes to eight classes. The candidates were classified according to classical virulence factors, genes associated with the cell surface, stress response genes, putative genes involved in DNA recombination and repair, genes required for transport, putative genes involved in metabolism, genes with regulatory roles and unknown genes.

5.1 Classical Virulence Genes:

In most STM studies, the majority of genes identified are not considered classical virulence determinants such as adhesions, invasions, and toxins. This is probably due in part to the evaluation of virulence of a human pathogen in an animal model. What this requires is an accurate reproduction of the natural route of infection which is not always reliable or available. Nevertheless, these types of genes are identified in STM screens validating the technique. Within the first category, insertions of the transposon were found in one classical virulence gene that corresponded to previously identified *cdtA* cytolethal distending toxin gene of *H. ducreyi* and a putative toxin-antitoxin cassette, *hicB*.

5.1.1 Cytolethal Distending Toxin A (*cdtA*):

In this study, an insertion was discovered in the cytolethal distending toxin protein A gene (*cdtA*). CdtA is a well characterized virulence factor that initiates eukaryote cell cycle block at G2 stage prior to mitosis. This toxin is composed of three proteins, CdtB, potentiates a cascade leading to cell cycle block, while, CdtA and CdtC which functions as dimeric subunits that bind to CdtB and deliver it to the mammalian cell interior. CdtB has DNase I-like activity that induces double strand breaks in host DNA. It has been found in other pathogens such as *E .coli, Campylobacter sp., Salmonella enterica server Typhi, Shigella dystenteria, Helicobacter sp* and *Actinobacillus actinomycetemcomitans.* Cytolethal distending toxin (CDT) inhibits cellular and humoral immunity via apoptosis of immune response cells. It generates necrosis of epithelial type cells involved in repair of lesions and thus the healing process is slower (126).

In the TDRM of *H. ducreyi* infection, *cdtA* mutant had moderate cytotoxicity on HeLa cells but the *cdtC* and *cdtB* mutants did not (78). In the TDRM, the *cdt* genes were transcribed in the rabbits demonstrated by RT-PCR. The ability of these mutants to colonize the skin or to cause experimental lesions compared to *H. ducreyi* 35000 was not noticeably different and formed at similar rates. When the *cdtC* mutant was co-injected with CDT the size and number of skin lesions increased. These results suggest that CDT may not be required for chancroid lesion formation but is important for infection (2, 134). Other genetic studies have indicated that to obtain an active cytotoxic effect, all three genes need to be expressed (109).

The identification of a CdtA mutant in this screen was a surprising discovery because it is a subunit of secreted toxin. Extracellular secreted toxins are thought to be complemented by other mutants in the inoculum. STM studies of other pathogens do not generally uncover "genuine" virulence factors like cytotoxins or effectors. However, an STM screen of Enterohemorrhagic *E. coli* O26:H$^-$ (EHEC) intestinal colonization of calves did in fact identify four mutants with insertions associated with genes encoding cytotoxins (148). Two of the mutants encoded for an enterohemolysin A and a secreted serine protease (148). Surprisingly, in a related study using STM to identify the O157:H7 genes involved in intestinal colonization of calves, none of the 79 characterized mutants encoded cytotoxins (41).

The STM screen for *H. ducreyi* was only tested once, and perhaps in a second screening with a pool of putative attenuated candidates, the *cdtA* mutant would not have the same phenotype. The CI calculated from an *in vitro* competition assay of this mutant did not show much of a deficiency in media. Also, the mutant was not actually assessed for attenuation of virulence via individual inoculation into the TDRM. Furthermore, it has been noted in some STM studies, when the attenuated phenotype of candidate mutants have been reassessed individually in the animal host or through another means of determining their virulence such as LD_{50}, the original assignment of attenuation was incorrect despite consistent lack of recovery from the mixed pools (41, 122). The three CDT genes are situated in an operon with *cdtA* being the first gene. It is highly likely that this knockout mutant represents one that contains a polar mutation and therefore displayed attenuation in the TDRM. Another possible explanation was that it represents a technical artifact of PCR amplification through competition with a different amplicon,

that this mutant was identified. Great care was taken to avoid such a false "positive" because each STM pool was tested in quadruplicate in two rabbits and only mutants that did not produce a PCR amplicon in all recovered tissues were considered attenuated candidates for further analysis. Further *in vivo* analysis of this mutant will help to determine its relative importance. The fact that bacterial virulence is a multifactorial process, it is possible that the interaction of *H. ducreyi* antigens or other cytotoxins and the human immune response play a role in the tissue destruction found in chancroid.

5.1.2 Hypervariable Intergenic Junction (*hicB*):

Another candidate for a classical virulence determinant was an insertion identified in a putative novel RNA-targeting toxin-antitoxin system (TAS). This site was first identified as an insertion into a major pilus gene cluster that correlated with colonization or pathogenicity (41, 88). Pili are hypothesized to contribute to pathogenesis by facilitating tissue-specific attachment to host cell. Upon sequencing the hypervariable left of the *hif-2* junction, genes that contained the *hicB* intergenic sequences, it was found that these conserved intergenic DNA were complex assemblies with evident frequent recombination and rearrangements. HicB was thought to be a "mobile" pathogenicity island but was not associated with a phage, transposon or insertion element. The Hic proteins may be involved in adhesion or function to block pilus-mediated attachment and were necessary for invasion of certain niches (114).

A more recent study, using bioinformatics and comparative genomic analysis revealed that it may be a Type II TAS because the region seems to be part of a two component system. HicB contains a partially degraded RNAse H fold and HicA has a

double-stranded RNA-binding domain. The RNAse H-like domain is generally fused to a DNA-binding domain (80). Type II TAS that are chromosomally located, in contrast to plasmid-encoded, enhance the physiology of the cell in response to external cues to induce reversible bacteriostasis or cell death. An example of this type of TAS, is the *relBE* locus that has been identified in *E. coli* K12. RelB sequesters the RelE toxin. RelE is an inhibitor of translation and induces bacteriostasis that can be reversed if the RelB antitoxin is produced. The *relBE* locus is activated in response to amino acid starvation and elevated levels of 3',5'guanosine bispyrophosphate. RelE cleaves mRNAs within the ribosome making them nonfunctional so that they cannot be processed further. This results in the accumulation of stalled ribosomes on damaged mRNAs and translation inhibition which in turn blocks cell cycle progression (29, 107). Most bacteria seem to contain many copies of chromosomally-located TAS suggesting that this is a wide-spread strategy for cell survival under stressful conditions which may pertain in particular to *H. ducreyi* (54). *H. ducreyi* must undergo many stressful circumstances in order to establish infection from enduring a change in optimal temperature to the limitation of nutrients within the host. In fact, in this STM screen, seven genes associated with stress response were identified in this pilot study.

5.2 Surface Molecules:

In the majority of STM studies, surface molecules represent 25-50% of the candidates (9). This underlies the importance that these molecules have in pathogenesis. Pathogens must come into contact with host cells and participate in critical processes for survival such as cell tropism, multiplication, stress adaptation, colonization and

dissemination. In this pilot *H. ducreyi* STM screen, about 15% of the candidates were surface molecules. Five genes are included in the second category of genes which are associated with the cell surface. Insertions were discovered in the *hgbA* hemoglobin receptor gene, *tdhA* ton-dependent heme receptor gene, *lpxB* lipid A disaccharide synthase, the gene coding for the large supernatant protein, *lspA1* and *lgtA*, a gene involved in cell wall biogenesis.

5.2.1 Hemoglobin Receptor A (*hgbA*):

Two important cell surface receptors for heme acquisition were identified that have been partially characterized previously by other researchers. The identification of these receptors reflect *H. ducreyi*'s dependence on the host to fulfill its heme requirement. An insertion in the previously characterized hemoglobin-binding gene, *hgbA* was discovered. HgbA is an outer-membrane protein that shares homology with other TonB-dependent receptor proteins, which are involved in heme or iron acquisition and transport (42).

An *hgbA* mutant is unable to bind or utilize hemoglobin as a source of heme (4). The *hgbA* mutant was shown to be attenuated in a host because the lesions produced in the rabbit model were much smaller compared to the wild type parent *H. ducreyi* 35000 (135). The *H. ducreyi* human experimental model of infection was tested with the *H. ducreyi hgbA* mutant FX504 on the upper arm of human subjects. Inoculation of the mutant did not result in pustule formation and was not recovered from biopsy specimens, even using a dose 10 times higher than the infective dose of the parent strain. These results suggest that hemoglobin may be an important source of heme or iron for *H.*

ducreyi to cause infection in humans (4). In the swine model of *H. ducreyi* infection, immunization with HgbA prevented chancroid. Histological sections of immunized animals lacked the typical features of chancroid and no viable *H. ducreyi* cells were recovered from inoculated sites. Anti-HgbA IgG blocked hemoglobin binding to the HgbA receptor. The identification of this gene within the STM screen asserts the importance of heme acquisition and suggests that the limitation of heme/iron acquisition of *H. ducreyi* could be used as a novel protection or vaccine strategy (1).

5.2.2 TonB Dependent Heme Receptor (*tdhA*):

Another outer membrane receptor for heme has been discovered in *H. ducreyi.* It was designated as *tdhA,* a ton-dependent heme receptor gene because it shares homology with other TonB-dependent receptor proteins. These proteins are involved in heme or iron acquisition and transport (140). There was an insertion of the transposon mapped to the *tdhA* gene. Most receptors for host iron-binding compounds belong to a family of outer membrane receptors designated as TonB dependent because they require the cytoplasmic membrane protein TonB for internalization of bound ligand. TonB with its accessory proteins, ExbB and ExbD, comprise a system for the transfer of energy from the cytoplasmic membrane to outer membrane receptors. Expression of the TdhA protein is regulated *in vitro* in heme-limiting conditions. However, in wild type, there was 50 to 100 times more HgbA then TdhA made. The ton system mutant had increased expression of the two proteins but had lower intracellular heme or iron levels. TdhA was also found to be widely distributed in *H. ducreyi* by Western blotting with peptide specific anti-sera. This suggests that it plays an essential role in these pathogens. The

poor expression of TdhA *in vitro* suggests that it may have a more important role to be expressed under certain conditions *in vivo* (140).

The TonB-dependent receptors have been evaluated for their role in virulence in humans. In conjunction with the two previously described receptors, a third receptor, TdX was also evaluated which shows homology to *yddB* in *E. coli.* The *yddB* mutant demonstrated decreased growth in rich media. *tdX* transcripts have been found in pustules of *H. ducreyi*-infected human volunteers. A double *tdX/tdhA* mutant (FX527) was found to be as virulent as its isogenic parent in the human experimental model of chancroid. This suggested that TdhA or TdX is not required for pustule formation in humans (75). However, in the human experimental model of chancroid lesions are not allowed to progress to the ulcerative stage of disease. This leaves open the possibility that expression of TdhA may contribute to the ulcerative stage.

The question that still remains to be answered is why TdhA could not compensate the HgbA mutant. *H. ducreyi* has been reported to have an extracellular and an intracellular life, in which case, the heme requirement must be met in both environments. In the extracellular environment after lysis of host erythrocytes by *H. ducreyi*, hemoglobin could be acquired via the HgbA receptor. In the intracellular milieu, hemoglobin could not serve as a heme source but heme synthesis in the cytoplasm of eukaryotic cells could fulfill the heme or iron requirement of *H. ducreyi* within the host through the TdhA receptor. Nonetheless, the fact that both of these receptors were identified in this STM screen demonstrates their significance in the biology of *H. ducreyi* infection within a host. The elucidation of heme acquisition mechanisms and their

subsequent exploitation for preventative therapies and vaccine strategies still remain to be of great importance.

5.2.3 Lipid Precursor (*lpxB*):

Surface molecules are important for the pathogen's ability to colonize mucosal surfaces. The first step in the colonization of host cells is to bypass the mucosal ciliary action through lipopolysaccaride (LPS). Actual LOS genes were not identified in this screen probably due to the route of inoculation used which circumvented this initial event. However, an insertion was found in a gene that mapped to *lpxB* which is a precursor of lipid that anchors LPS to the outermembrane. The *H. ducreyi lpxB* has 96% homology to the *lpxB* gene that was first characterized in *H. influenzae.* LpxB catalyzes the formation of lipid A disaccharide from UDP-2,3-diacylglucosamine and 2,3-diacylglucosamine-1-phosphate. In *Legionella pneumophila,* differential regulation of *lpxA* or *lpxB* paralogues in response to changing environmental conditions such as nutrient deprivation and hypotonic stress may allow the pathogen to adapt its lipid A structure (5). Adaptation of LPS in response to altered conditions would serve as a crucial mechanism for *H. ducreyi* survival as it encounters nutrient limitations and host immune cells.

5.2.4 Lipooligosaccharide N-acetylglucosamine Glycosyltransferase (*lgtA*):

Genes encoding cell wall elements involved in such activities as peptidoglycan biosynthesis or modification are found in a number of STM studies. In this study, a gene was identified that maps in an *H. ducreyi* gene whose product is involved in cell wall

biogenesis of the outer membrane. Lipooligosaccharide N-acetylglucosamine glycosyltransferase, LgtA, transfers a sugar from UDP-glucose to a range of substrates such as cellulose, acetyl-galactosamine, GDP-mannose, dolichol phosphate and teichoic acids. An *H. ducreyi lgtA* mutant had LPS that lacked N-acetylglucosamine distal sugars (138). It has been proposed that the LOS evades the human immune response through molecular mimicry because the structure of the oligosaccharide section of LOS looks a lot like the structure of human paraglobosides (81). A stronger inflammatory response and the formation of ulcers in rabbits is observed after injection with *H. ducreyi* LOS compared to injection with *E. coli* LPS (84). In many gram-negative pathogens, LOS is an important virulence determinant and changes in LOS biosynthesis may result in decreased resistance to the bactericidal activity of human serum. Serum resistance of *H. ducreyi* may be attributed to its LOS content and structure. Characterization of the composition of LOS would contribute to the understanding of the susceptibility of *H. ducreyi* to the bactericidal action of normal human serum.

5.2.5 Large Supernatant Protein (*lspA1*):

Adherence of bacterial cells to host epithelium would be the next step in host cell colonization. A mutation in the large supernatant protein, *lspA1* was identified in this screen which has been previously analyzed. LspA1 has haemagglutination activity and in the *Yersinia* homolog, *FhaB*, is involved in heme utilization or adhesion (127). Expression of LspA1 and LspA2 were required for full virulence in the human model of *H. ducreyi* infection (65). An *H. ducreyi* mutant deficient in both LspA1 and LspA2 was significantly less virulent than *H. ducreyi* 35000 in the temperature-dependent animal

model (149). To date, there is evidence that it is required to form lesions because avirulent *H. ducreyi* strains tested in the TDRM did not produce *lspA1* or *lspA2* genes *in vitro* and recombinant LspB did not form lesions when injected into the TDRM (150). This study validates the importance of this gene in the pathogenesis of *H. ducreyi* infection within a host.

5.3 Stress Response Genes:

Members of the third group of genes identified encode proteins involved in stress response. Four stress related genes with insertions were highlighted, *dnaK, uspA, sodA* and *ptrA.* These types of genes participate in essential processes of adaptation to survive and are discovered in many STM studies.

5.3.1 Heatshock Chaperone Protein (*dnaK*):

A gene encoding a heatshock protein, DnaK was identified in this screen. Heat shock proteins stabilize essential and virulence-related proteins in bacterial pathogens during exposure to environmental stress by acting as molecular chaperones (51). The heat shock response in prokaryotes involves the transient expression of Hsp60 (GroEL), Hsp70 (DnaK) and other heat shock proteins regulated both positively and negatively. DnaK is part of the primary stress-sensing and transduction system where it modulates the induction of the heat shock response. DnaK is probably best described in *E. coli* K12. The proposed model for the Hsp70(DnaK) chaperone cycle is as follows: upon ATP binding, the ATPase and the substrate binding domains of DnaK are docked and the substrate binding cavity is opened The ATP-bound DnaK chaperone has low affinity and

fast exchange rates for its substrates. A cochaperone (JDP) with bound substrate interacts with its DnaK cognate partner which induces a conformational change in the ATPase domain. This stimulates ATP hydrolysis and the closing of the substrate binding cavity. In this state, DnaK exhibits high affinity and low exchange rates for substrates. The nucleotide exchange factor, GrpE aids the release of ADP via its interaction with DnaK ATPase. The binding of ATP changes the conformational state of DnaK and releases the substrate, thus resetting the DnaK cycle (51).

In an STM screen of *Actinobacillus pleuropneumoniae* infection of the respiratory tract of swine, the cochaperone DnaJ was identified along with another heatshock protein, HtpG (123). For *H. ducreyi*, it is likely that heat shock proteins play a critical role in its survival within a host due to its temperature sensitivity. It grows best *in vitro* at 33-35°C and on the (lower temperature) surface of a host which may enhance its transmission during sexual contact. DnaK may be a negative modulator of heatshock gene expression through the interaction with the sigma-32 homolog because when DnaK was overexpressed in *H. ducreyi,* GroEL expression decreased (105). A 58.5 kDa Heatshock protein (Hsp) found in *H. ducreyi* suggested that it was involved in the attachment of bacteria to host cells and to each other. It was associated with the bacterial surface visualized by immunostaining and adhered to HEp-2 cells (48). RNA and protein analysis show that the *groE* operon was highly expressed in exponential and stationary phase. In addition, the analysis revealed that the *groE* operon was also highly expressed when *H. ducreyi* was grown at temperatures higher than 33°C (106). Stationary phase is thought to resemble the conditions that most bacteria encounter for much of their existence *in vivo*. When the DnaK STM mutant was tested in the TDRM individually,

the mutant demonstrated only partial attenuation in virulence. Much of this process still needs to be elucidated but the data suggests that the transcriptional regulation of the heatshock operon is a coordinated response to protect and promote the survival of *H. ducreyi* within a host at an elevated temperature.

5.3.2 Universal Stress Protein A (*uspA*):

A mutation was mapped to the universal stress protein A (*uspA)* gene found in *A. pleuropneumoniae.* The *H. ducreyi* UspA showed 96% homology to the one found in *A. pleuropneumoniae*. It is a small cytoplasmic bacterial protein whose expression is upregulated when the cell is exposed to stress agents such as DNA damaging agents and those that arrest cell growth. UspA enhances the cell survival during prolonged exposure to nutrient starvation and heat shock (128). The activity of this protein is to maintain general stress endurance and further investigation will be required to determine the relative importance of this protein in the pathogenesis of *H. ducreyi.*

5.3.3 Superoxide Dismutase A (*sodA*):

A gene that encoded the characterized *H. ducreyi sodA* gene was identified in this study. This manganese superoxide dismutase catalyzes the conversion of superoxide radicals to hydrogen peroxide and molecular oxygen (133). These enzymes play a crucial role in the maintenance of cell and DNA integrity during aerobic metabolism by eliminating reactive oxygen species that induce lipid damage and the oxidative inactivation of essential enzymes. The *H. ducreyi* SodA can complement the aerobic growth defect of an *E. coli sodA sodB* double mutant in minimal media (118). For the

STM mutant, the *in vitro* CI = 0.708, demonstrated that it was only slightly deficient in growth compared to the wild type parent strain. This suggests that SodA has a more prominent role *in vivo* because it was identified in the STM screen, which still needs to be further characterized.

5.3.4 Protease III (*ptrA*):

A protease III gene, *ptrA*, was found to have 38% amino acid identity and 59% homology with the one from *E. coli.* PtrA is a secreted periplasmic Zn-dependent peptidase that is involved in posttranslational modification, protein turnover and chaperones. The *E. coli ptrA* mutant did not show any phenotypic alterations in the bacteria and overexpression of protease III does not appear to have any effects on cell viability. Interestingly though, the *E. coli ptrA* gene is located between *recC* and *recB* which are part of an operon important for DNA recombination and repair. It has been suggested that it is involved in degradation of unwanted proteins and is functionally compartmentalized so it would be available at only certain times and places (40). For *H. ducreyi*, it seems that it may an important protein required for survival within an animal host, perhaps by degrading host proteins as they enter the extracellular space but further investigation must be done to obtain it role in *H. ducreyi* pathogenesis.

5.4 DNA Recombination and Repair Genes:

The fourth group includes genes that are related to the previous category of stress response genes. Three mutants had insertions in genes that show similarity to those involved in DNA recombination and repair. Disruption of these genes would impair the

fitness of *H. ducreyi* in an animal host. Specifically, three genes involved in mutation avoidance such as *recR*, *radA*, and *mutS* were identified.

5.4.1 Recombination R (*recR*):

The gene identified as *recR*, displayed 90% homology to the one found in *H. influenzae*. RecR is involved in the recombinational process of DNA repair. It is part of the RecOR complex that displace the ssDNA binding protein from ssDNA (61). Mutations in recR delay induction of the SOS response in *E. coli* and these mutants are sensitive to UV irradiation and mitomycin C (71).

5.4.2 DNA-dependent ATPase (*radA*):

A gene encoding *radA* was similar to an ortholog of the *radA* gene of *E. coli*, which showed 89% homology and 76% amino acid identity. RadA is a DNA repair protein that stabilizes the strand-invasion intermediate during DNA repair after exposure to mutagenic agents. In *E. coli, radA* mutants showed a moderate decrease in survival after UV, hydrogen peroxide, hydroxyurea, phleomycin, methyl methanesulfanote and X-irradiation exposure. The mutant was deficient in post-synaptic DNA processing (16).

5.4.3 Mismatch (*mutS*):

A mutation that mapped to the *mutS* gene found in *H. influenzae* with 82% homology and 71% amino acid identity. MutS is involved in mismatch recognition during the DNA repair process. In *H. influenzae,* it binds ssDNA during the exonucleolytic step of DNA repair (67). During the infection, pathogenic bacteria may

encounter a variety of oxygen radicals produced by macrophages and neutrophils. These oxygen species are likely to induce mutations in the DNA. These genes identified suggest that DNA repair is an important mechanism for *H. ducreyi* survival *in vivo* and future work is required to dissect their complete role in pathogenesis.

5.5 Transport System Genes:

The fifth class contains mutants showing insertions in three genes involved in transport. Interestingly, there were no ABC transporters identified in this screen which usually make up to 30% of the genes discovered in most STM studies. ABC transport systems are involved in the uptake or efflux of a diverse array of macromolecules across the cytoplasmic membrane of bacteria and eukaryotes. Many of the ABC transport proteins identified in STM studies are involved in nutrient acquisition.

5.5.1 Outermembrane Lipoprotein Carrier (*lolA*):

An insertion mapped to a gene with similar function as an outermembrane lipoprotein carrier, *lolA* from *H. influenzae*. Lipoproteins are usually anchored to the outer membrane through N-terminal lipids. Lipid modification and processing to mature lipoproteins occur on the periplasmic surface of the inner membrane.

In *E. coli*, LolA participates with LolB in the incorporation of lipoprotein into the outermembrane by acting as a periplasmic chaperone (89). LolA is thought to undergo a conformational change when the lipoprotein is bound and is hydrophilic so that it can travel through the periplasm. An ATP binding cassette (ABC) transporter, LolCDE, in the inner membrane releases outer membrane-specific lipoproteins. LolCDE catalyzes

the extrusion of lipoproteins anchored to the outer leaflet of the inner membrane (139). Lol proteins are essential for *E. coli* growth and are widely conserved in Gram-negative bacteria (89, 93). Bacterial cell wall lipoproteins are proinflammatory which make them important for pathogenesis. As for *H. ducreyi*, a peptidoglycan-associated lipoprotein (PAL) has been identified and shown to be required for virulence in the human model of *H. ducreyi* infection (46). Mutations of this gene disrupted outer membrane integrity and increased the strain's hypersensitivity to antibiotics (46). For *H. ducreyi* the importance of ATP-dependent lipoprotein release *in vivo* has been revealed and again demonstrates that structural integrity of the membrane of the pathogen is necessary for survival within a host through the induction and evasion of the innate and adaptive immune responses.

5.5.2 Sodium/Proline Symporter (*putP*):

A gene was found to encode the proline permease component of a sodium/proline symporter, *putP*. There was 76% homology to the one from *H. influenzae*. PutP catalyzes the Na^+ coupled uptake of proline with a 1:1 stoichiometry (68). To energize proline transport, the electrochemical Na^+ gradient is generated by the action of Na^+/ H^+ antiporter(s). Subsequent to the uptake process, proline is oxidized via pyrroline 5-carboxylate to L-glutamate. In *Staphylococcus aureus*, this high affinity transporter contributes to its *in vivo* survival. Insertional inactivation of the *putP* gene in *S. aureus* demonstrated reduced virulence in experimental endocarditis (15). Thus, PutP may represent a new target for the development of new drugs against pathogens and additional characterization of the genes is necessary to elucidate its role in *H. ducreyi* pathogenesis.

5.5.3 Multidrug Resistant Protein (*norM*):

Insertions were also found in genes that showed similarity to *norM*, a multidrug resistance protein. An orthologue gene *norM* from *Vibrio parahaemolyticus* was identified in the screen. There was a 63% homology to this gene. NorM is an efflux pump that is part of the MATE family of Na(+)/drug antiporters and in *Vibrio* it confers resistance to norfloxacin, ciprofloxacin, doxorubicin, daunomycin, deoxycholate, DAPI and ethidium bromide. A *V. parahaemolyticus norM* mutant has been shown to be sensitive to fluoroquinolones (91). In *H. influenzae*, the gene responsible for norfloxacin resistance, *hmrM* was cloned and showed sequence similarity to NorM. There were some substrate differences observed and HmrH showed the highest resistance to acriflavine and ethidium bromide while NorM showed the highest resistance to doxorubicin, daunomycin and ethidium bromide. Both showed a similar resistance to norfloxacin (153). To date, *H. ducreyi* has not shown resistance to norfloxacin.

Generally, STM screens identify a larger number of genes associated with transport than the three found in this *H. ducreyi* screen. The biology of *H. ducreyi* itself could be the limiting factor of identifying a low number of ABC transporters. It is a fairly inert pathogen with a moderately small compact genome of 1.7 Mb. The total number of transporters within the genome is 88, with 30.7% of those being of the ATP-dependent type. Pathogens with compact genomes are known to have ABC transporters with dual functions to compensate and acquire all the essential nutrients from their environment. Another explanation for the low frequency of genes involved in transport being identified is due to the modest number of mutants screened and further experiments

testing more mutants in the rabbit model may be more fruitful in the discovery of important transport genes involved in *H. ducreyi* pathogenesis.

5.6 Metabolism Genes:

The ability to adapt to the host environment is a major component of pathogenicity. A few genes whose products are involved in metabolic pathways and nutrient uptake have been identified following the negative selection *in vivo*. The sixth group includes insertions found in genes coding for polypeptides that show similarity to ones involved in metabolic functions.

5.6.1 Fumarate Reductase Flavoprotein (*frdA*):

A gene coding for a protein similar to a fumarate reductase flavoprotein, *frdA,* was identified. It showed 73% identity to the one in *E. coli*. FrdA reduces fumarate to succinate in anaerobic bacterial respiration. Fumarate reductase is important for both energy metabolism and succinate biosynthesis. In *Helicobacter pylori*, three antihelmintics, oxantel, thiabendazole, and morantel were shown to inhibit its activity, stopping cell growth leading to cell death. This suggests that FrdA could be used as a therapeutic target for the eradication of *H. pylori* (50). An *frd* mutant displayed a threefold decrease in cell mass yield compared to wild type when grown with glucose. FrdA has an important role in heme-dependent growth stimulation of *Bacteroides fragilis* (14). Analysis of the *H. ducreyi* genome indicates that is has an incomplete, reductive TCA cycle which is the case for many bacteria (59). As for fermentation, *H. ducreyi* contains the genes necessary to convert pyruvate to lactate, formate and acetate but not to

ethanol (24). It has been proposed that the incomplete reductive type of cycle functions in carbon assimilation and the generation of precursor metabolites for biosynthesis (59).

5.6.2 2-oxoglutarate Dehydrogenase (*sucA*):

A gene involved in anaerobic respiration, 2-oxoglutarate dehydrogenase (*sucA*) which is an alpha-ketoglutarate decarboxylase was identified that had 86% homology with the gene from *H. influenzae*. SucA is involved in the citrate cycle, lysine degradation and tryptophan metabolism. It produces succinic semialdehyde that is part of the alternative pathway from alpha-ketoglutarate to succinate which is essential for normal growth. *Mycobacterium tuberculosis* encodes *kgd* and produces succinic semialdehyde which forms an alternate pathway from α-ketoglutarate to succinate (142). *H. ducreyi* has not shown to be able to grow anaerobically but it has been reported that it has an extracellular and an intracellular life style. This dual life style would necessitate that the heme requirement be met in both environments and that genes involved in anaerobic respiration be expressed. An interesting fact is that Kgd is not found in humans and may represent a potential target for antimicrobial therapy.

5.6.3 Cytochrome C-type Biogenesis (*ccmH*):

An insertion was mapped to a for the cytochrome c-type biogenesis, *ccmH*, from *E. coli*. Cytochrome c maturation protein H (ccmH) is a membrane-anchored thiol-oxidoreductase that is thought to be involved in the reduction of apocytochrome *c*, a prerequisite for covalent heme attachment. It plays a crucial role in the maturation of cytochromes in the periplasm of Gram-negative bacteria (39). Interestingly, the gene in

H. ducreyi contains an authentic frame shift that resulted in a stop codon. Thus this gene is described as a pseudogene and is probably nonfunctional which makes it a surprising find. This discovery could represent a technical artifact of PCR amplification or perhaps the transposon inserted in multiple loci and this was the gene that was sequenced. Further analyses of this mutant will help determine its relative importance. Nevertheless, cell wall biosynthesis genes are important to the survival of *H. ducreyi* in vivo.

5.6.4 Aspartyl-tRNA Synthetase (*aspS*) and Prolyl-tRNA Synthetase (*proS*):

Two insertions were mapped to genes similar to *aspS* and *proS* which belong to tRNA synthetases. Theses enzymes are crucial in the biochemical process in every cell to provide the correct amino acid substrates for ribosomal translation. It was found that the *aspS* gene displayed 83% homology to the *E. coli aspS* gene. AspS is an aspartyl-tRNA synthetase that codes for AspRS that is a required intermediate in protein synthesis. It has the ability to aminoacylate two different tRNA isoreceptors involved in alanine and aspartate metabolism (60). The *proS* gene showed 91% homology to the prolyl-tRNA synthetase gene from *H. influenzae.* ProS catalyzes the formation of prolyl-tRNA from proline and tRNA via the attachment of proline to the 3'OH group of ribose on the appropriate tRNA (63). These data suggest that *in vivo*, *H. ducreyi* may require increased alanine, aspartate and proline. These semi-essential amino acids may be in limited supply in host tissues and thus these biosynthetic pathways are required for infection.

5.7 Regulatory Genes:

Another group of genes were classified as regulatory genes which are important for the control of gene expression. In the multifactorial pathogenesis process, it is a crucial activity to turn on genes at the right time for maintaining and developing the infection.

5.7.1 Fumarate/nitrate Reduction Transcription Regulator (*hlyX*):

A gene encoding *hlyX*, a fumarate/nitrate reduction (Fnr-like) transcriptional regulator from *H. influenzae* was identified that showed 90% homology. HlyX is a global transcription factor that controls the expression of over 100 target genes in response to anoxia. The *E. coli* FNR homolog regulates hemolytic activity by using fumarate and nitrate as reducible substrates for anaerobic growth (31). Anaerobic growth in *H. ducreyi* has not been reported but it is thought to have an intracellular life style. The regulation of these genes may serve as an important factor during *H. ducreyi* infection.

5.7.2 Transcription Cleavage Factor (*greA*):

A gene encoding a transcription elongation factor, *greA* was identified that showed 92% homology to the one from *H. influenzae.* GreA is necessary for efficient RNA polymerase transcription past template-encoded arresting sites that trap elongating RNA polymerases that pass through. In *E. coli*, GreA induces cleavage of nascent transcript in ternary complexes of RNA polymerase by suppression of elongation arrest and by enhancement of transcription fidelity (21). *E. coli* cells lacking GreA and GreB

(another transcriptional elongation factor) are hypersensitive to temperature changes (100). GreA was found to be essential for *Yersinia pestis* virulence in a subcutaneously infected mouse model (45). Efficient RNA polymerase transcription in *H. ducreyi* has not been looked at before but the identification of this gene implicates the significance of this process *in vivo*. The changes that *H. ducreyi* must undergo in response to environmental signals may directly contribute to the control of its virulence genes.

5.8 Genes with Hypothetical Functions:

Finally, the last category of genes identified in STM studies, are those that have some sequence similarity with genes of known function. In some STM studies this represents 50% of the genes identified. This number also includes genes that have no significant similarity with genes of known function or organism-specific sequences with no known homologues. Only two genes with hypothetical function were identified in this study. One insertion was found in a gene encoding for a hypothetical aminotransferase, HD0464. Gene HD0464 shows some putative homology to selenocysteine lyase/cysteine desulfurase that may be involved in posttranslational modification and protein turnover. HD1268 had an insertion in a hypothetical gene that had a predicted esterase/lipase function.

5.9 Comparison of *in vivo* Genes Identified with STM and SCOTS:

In a complementary study, *H. ducreyi* genes that were expressed during human infection were identified (13). They used Selective Capture of Transcribed Sequences (SCOTS) with RNA isolated from pustules obtained from volunteers infected with *H.*

ducreyi. The STM results were similar to those reported in the SCOTS-based study and identified genes that belonged into the same functional categories. A total of 531 genes were identified and 139 of those genes were found in two volunteers. Many genes were identified that had roles in metabolism, DNA replication/repair, stress response and hypothetical functions. Only 10 out of 139 genes that were associated with transport were identified. Interestingly, they found genes that were involved in anaerobic respiration as well. Three genes that were discovered using SCOTS were also identified by this STM study including *radA*, *mutS* and *lspA1* which highlight the significant role that these genes play *in vivo* (13). The striking similarities of the results from the SCOTS experiment corroborate the different types of genes identified from this STM study and validate the TDRM as an animal model for chancroid.

5.10 Future Prospects - Characterizing the Candidates:

Further analyses of the three candidates, (the *tdhA*, *dnaK*, and *hicB* mutants) include complementation of the attenuated virulence phenotype that was observed in the rabbit model. The genes of interest (*tdhA*, *dnaK*, and *hicB*) would be cloned into an expression vector and electroporated into the mutants to restore the attenuated phenotype. These mutants would then be tested in the rabbit model to assess their virulence. Experiments to confirm that these genes are responsible for the virulence attenuated phenotype would be to reconstruct the mutants using in-frame insertional deactivation of the gene with a kanamycin cassette from a vector pUC18K and through homologous recombination create the knockouts. The kanamycin cassette from this vector has been designed to create nonpolar mutations (86). This would provide information as to

whether there was a polar effect of the transposon on the downstream genes. Further research could also be performed to define the precise role of the gene product by determining their cellular location, biochemical activity and their interacting partners and structure.

The other virulence attenuated candidates would also be analyzed through individual inoculation into the rabbit model to assess their degree of attenuation. These analyses would confirm the attenuated phenotypes of the candidates and present new and interesting *in vivo* virulence determinants of *H. ducreyi* infection for further functional characterization.

5.11 Summary Statements:

In conclusion, this study provides the results of the first large-scale screening of STM mutants of *H. ducreyi* and the identification of virulence genes in a rabbit infection model. Most of the work performed in this study was to establish a working STM model of *H. ducreyi* infection. The initial experiments justified this study as a proof of principle for the STM approach for the *in vivo* identification of virulence determinants in *H. ducreyi*. Since the ground work has been completed, the STM approach can now be used to screen many more *H. ducreyi* mutants through the rabbit model. This STM screen was not saturating and there are likely many more factors to be discovered that contribute to the ability of *H. ducreyi* to survive *in vivo*.

The STM study resulted in the identification of known virulence genes and a large number of genes that were not previously known to play a role in the virulence of *H. ducreyi*. The identification of known virulence determinants such as the cytolethal

distending toxin, and genes involved in cell surface (*lspA1*) or cell wall biogenesis (*lgtA* and *lpxB*) validated the STM approach in *H. ducreyi*. Two genes important in heme uptake (*tdhA* and *hgbA*) were identified in this screen demonstrating the unique biology of *H. ducreyi* and the significance of iron/heme acquisition within a host. The identification of genes such as heat shock proteins (*dnaK* and *uspA*) and free radical scavengers (*sodA*) along with genes involved in DNA repair (*mutS*, *radA*, *recR*), emphasized the importance that stress response genes of a pathogen play during infection of host. Only three genes associated with transport were identified which indicates that the screen was not exhaustive because it is likely that more transport genes contribute to the pathogenesis of *H. ducreyi* due to their critical roles in nutrient acquisition. Five genes were discovered to be involved in metabolism which emphasized the significance that these factors contributed to the fitness of *H. ducreyi* within a host. In particular, two of these genes (*frdA* and *sucA*) may contribute to the pathogen's ability to grow under anaerobic conditions that have not been thoroughly analyzed before. The identification of two regulatory genes (*hlyX* and *greA*) demonstrates that virulence genes respond to signals that belong to well-characterized global regulatory networks. As in other STM studies, genes with unknown and hypothetical functions were identified in this screen. One gene, *hicB*, in particular is suggested to have RNA toxin-antitoxin characteristics. Many of these determinants will need to be further addressed to define their precise host-pathogen interaction role. The discovery of a variety of different genes, some that were not previously implicated in pathogenesis, suggests that there are many future paths of study towards elucidation of virulence mechanisms in *H. ducreyi*.

CHAPTER 6: REFERENCES

1. **Afonina, G., I. Leduc, I. Nepluev, C. Jeter, P. Routh, G. Almond, P. E. Orndorff, M. Hobbs, and C. Elkins.** 2006. Immunization with the *Haemophilus ducreyi* hemoglobin receptor HgbA protects against infection in the swine model of chancroid. Infect Immun **74:**2224-32.
2. **Ahmed, H. J., L. A. Svensson, L. D. Cope, J. L. Latimer, E. J. Hansen, K. Ahlman, J. Bayat-Turk, D. Klamer, and T. Lagergard.** 2001. Prevalence of cdtABC genes encoding cytolethal distending toxin among *Haemophilus ducreyi* and Actinobacillus actinomycetemcomitans strains. J Med Microbiol **50:**860-4.
3. **Ahn, S. J., Z. T. Wen, and R. A. Burne.** 2007. Effects of oxygen on virulence traits of *Streptococcus mutans.* J Bacteriol **189:**8519-27.
4. **Al-Tawfiq, J. A., K. R. Fortney, B. P. Katz, A. F. Hood, C. Elkins, and S. M. Spinola.** 2000. An isogenic hemoglobin receptor-deficient mutant of *Haemophilus ducreyi* is attenuated in the human model of experimental infection. J Infect Dis **181:**1049-54.
5. **Albers, U., A. Tiaden, T. Spirig, D. Al Alam, S. M. Goyert, S. C. Gangloff, and H. Hilbi.** 2007. Expression of *Legionella pneumophila* paralogous lipid A biosynthesis genes under different growth conditions. Microbiology **153:**3817-29.
6. **Albritton, W. L.** 1989. Biology of *Haemophilus ducreyi.* Microbiol Rev **53:**377-89.
7. **Ambrosi, C., F. Tiburzi, F. Imperi, L. Putignani, and P. Visca.** 2005. Involvement of AlgQ in transcriptional regulation of pyoverdine genes in *Pseudomonas aeruginosa* PAO1. J Bacteriol **187:**5097-107.
8. **Atkins, T., R. Prior, K. Mack, P. Russell, M. Nelson, J. Prior, J. Ellis, P. C. Oyston, G. Dougan, and R. W. Titball.** 2002. Characterisation of an acapsular mutant of *Burkholderia pseudomallei* identified by signature tagged mutagenesis. J Med Microbiol **51:**539-47.
9. **Autret, N., and A. Charbit.** 2005. Lessons from signature-tagged mutagenesis on the infectious mechanisms of pathogenic bacteria. FEMS Microbiol Rev **29:**703-17.
10. **Autret, N., I. Dubail, P. Trieu-Cuot, P. Berche, and A. Charbit.** 2001. Identification of new genes involved in the virulence of *Listeria monocytogenes* by signature-tagged transposon mutagenesis. Infect Immun **69:**2054-65.
11. **Badger, J. L., C. A. Wass, S. J. Weissman, and K. S. Kim.** 2000. Application of signature-tagged mutagenesis for identification of *Escherichia coli* K1 genes that contribute to invasion of human brain microvascular endothelial cells. Infect Immun **68:**5056-61.
12. **Bauer, B. A., M. K. Stevens, and E. J. Hansen.** 1998. Involvement of the *Haemophilus ducreyi* gmhA gene product in lipooligosaccharide expression and virulence. Infect Immun **66:**4290-8.
13. **Bauer, M. E., K. R. Fortney, A. Harrison, D. M. Janowicz, R. S. Munson, Jr., and S. M. Spinola.** 2008. Identification of *Haemophilus ducreyi* genes expressed during human infection. Microbiology **154:**1152-60.

14. **Baughn, A. D., and M. H. Malamy.** 2003. The essential role of fumarate reductase in haem-dependent growth stimulation of *Bacteroides fragilis*. Microbiology **149:**1551-8.
15. **Bayer, A. S., S. N. Coulter, C. K. Stover, and W. R. Schwan.** 1999. Impact of the high-affinity proline permease gene (putP) on the virulence of *Staphylococcus aureus* in experimental endocarditis. Infect Immun **67:**740-4.
16. **Beam, C. E., C. J. Saveson, and S. T. Lovett.** 2002. Role for radA/sms in recombination intermediate processing in *Escherichia coli*. J Bacteriol **184:**6836-44.
17. **Behari, J., L. Stagon, and S. B. Calderwood.** 2001. pepA, a gene mediating pH regulation of virulence genes in *Vibrio cholerae*. J Bacteriol **183:**178-88.
18. **Behets, F. M., J. Andriamiadana, D. Randrianasolo, R. Randriamanga, D. Rasamilalao, C. Y. Chen, J. B. Weiss, S. A. Morse, G. Dallabetta, and M. S. Cohen.** 1999. Chancroid, primary syphilis, genital herpes, and lymphogranuloma venereum in Antananarivo, Madagascar. J Infect Dis **180:**1382-5.
19. **Behets, F. M., A. R. Brathwaite, T. Hylton-Kong, C. Y. Chen, I. Hoffman, J. B. Weiss, S. A. Morse, G. Dallabetta, M. S. Cohen, and J. P. Figueroa.** 1999. Genital ulcers: etiology, clinical diagnosis, and associated human immunodeficiency virus infection in Kingston, Jamaica. Clin Infect Dis **28:**1086-90.
20. **Bong, C. T., R. E. Throm, K. R. Fortney, B. P. Katz, A. F. Hood, C. Elkins, and S. M. Spinola.** 2001. DsrA-deficient mutant of *Haemophilus ducreyi* is impaired in its ability to infect human volunteers. Infect Immun **69:**1488-91.
21. **Borukhov, S., A. Polyakov, V. Nikiforov, and A. Goldfarb.** 1992. GreA protein: a transcription elongation factor from *Escherichia coli*. Proc Natl Acad Sci U S A **89:**8899-902.
22. **Brown, J. S., A. Aufauvre-Brown, J. Brown, J. M. Jennings, H. Arst, Jr., and D. W. Holden.** 2000. Signature-tagged and directed mutagenesis identify PABA synthetase as essential for *Aspergillus fumigatus* pathogenicity. Mol Microbiol **36:**1371-80.
23. **Camacho, L. R., D. Ensergueix, E. Perez, B. Gicquel, and C. Guilhot.** 1999. Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. Mol Microbiol **34:**257-67.
24. **Challacombe, J. F., A. J. Duncan, T. S. Brettin, D. Bruce, O. Chertkov, J. C. Detter, C. S. Han, M. Misra, P. Richardson, R. Tapia, N. Thayer, G. Xie, and T. J. Inzana.** 2007. Complete genome sequence of *Haemophilus somnus* (*Histophilus somni*) strain 129Pt and comparison to *Haemophilus ducreyi* 35000HP and *Haemophilus influenzae* Rd. J Bacteriol **189:**1890-8.
25. **Chen, C. Y., R. C. Ballard, C. M. Beck-Sague, Y. Dangor, F. Radebe, S. Schmid, J. B. Weiss, V. Tshabalala, G. Fehler, Y. Htun, and S. A. Morse.** 2000. Human immunodeficiency virus infection and genital ulcer disease in South Africa: the herpetic connection. Sex Transm Dis **27:**21-9.
26. **Chen, C. Y., K. J. Mertz, S. M. Spinola, and S. A. Morse.** 1997. Comparison of enzyme immunoassays for antibodies to *Haemophilus ducreyi* in a community outbreak of chancroid in the United States. J Infect Dis **175:**1390-5.

27. **Chiang, S. L., and J. J. Mekalanos.** 1998. Use of signature-tagged transposon mutagenesis to identify *Vibrio cholerae* genes critical for colonization. Mol Microbiol **27:**797-805.
28. **Chiang, S. L., J. J. Mekalanos, and D. W. Holden.** 1999. In vivo genetic analysis of bacterial virulence. Annu Rev Microbiol **53:**129-54.
29. **Christensen, S. K., and K. Gerdes.** 2003. RelE toxins from bacteria and Archaea cleave mRNAs on translating ribosomes, which are rescued by tmRNA. Mol Microbiol **48:**1389-400.
30. **Coico, R.** 2001. Gram staining. Curr Protoc Immunol **Appendix 3:**Appendix 3O.
31. **Cole, S. T., C. Condon, B. D. Lemire, and J. H. Weiner.** 1985. Molecular biology, biochemistry and bioenergetics of fumarate reductase, a complex membrane-bound iron-sulfur flavoenzyme of *Escherichia coli*. Biochim Biophys Acta **811:**381-403.
32. **Cope, L. D., S. Lumbley, J. L. Latimer, J. Klesney-Tait, M. K. Stevens, L. S. Johnson, M. Purven, R. S. Munson, Jr., T. Lagergard, J. D. Radolf, and E. J. Hansen.** 1997. A diffusible cytotoxin of *Haemophilus ducreyi*. Proc Natl Acad Sci U S A **94:**4056-61.
33. **Cortes-Bratti, X., E. Chaves-Olarte, T. Lagergard, and M. Thelestam.** 1999. The cytolethal distending toxin from the chancroid bacterium *Haemophilus ducreyi* induces cell-cycle arrest in the G2 phase. J Clin Invest **103:**107-15.
34. **Courvalin, P., and C. Carlier.** 1987. Tn1545: a conjugative shuttle transposon. Mol Gen Genet **206:**259-64.
35. **Darwin, A. J., and V. L. Miller.** 1999. Identification of *Yersinia enterocolitica* genes affecting survival in an animal host using signature-tagged transposon mutagenesis. Mol Microbiol **32:**51-62.
36. **De Ley, J., W. Mannheim, R. Mutters, K. Piechulla, R. Tytgat, P. Segers, M. Bisgaard, W. Frederiksen, K. H. Hinz, and M. Vanhoucke.** 1990. Inter- and intrafamilial similarities of rRNA cistrons of the *Pasteurellaceae*. Int J Syst Bacteriol **40:**126-37.
37. **Desjardins, M., L. G. Filion, S. Robertson, and D. W. Cameron.** 1995. Inducible immunity with a pilus preparation booster vaccination in an animal model of *Haemophilus ducreyi* infection and disease. Infect Immun **63:**2012-20.
38. **Desjardins, M., L. G. Filion, S. Robertson, L. Kobylinski, and D. W. Cameron.** 1996. Evaluation of humoral and cell-mediated inducible immunity to *Haemophilus ducreyi* in an animal model of chancroid. Infect Immun **64:**1778-88.
39. **Di Matteo, A., S. Gianni, M. E. Schinina, A. Giorgi, F. Altieri, N. Calosci, M. Brunori, and C. Travaglini-Allocatelli.** 2007. A strategic protein in cytochrome c maturation: three-dimensional structure of CcmH and binding to apocytochrome c. J Biol Chem **282:**27012-9.
40. **Dykstra, C. C., and S. R. Kushner.** 1985. Physical characterization of the cloned protease III gene from *Escherichia coli* K-12. J Bacteriol **163:**1055-9.
41. **Dziva, F., P. M. van Diemen, M. P. Stevens, A. J. Smith, and T. S. Wallis.** 2004. Identification of *Escherichia coli* O157 : H7 genes influencing colonization of the bovine gastrointestinal tract using signature-tagged mutagenesis. Microbiology **150:**3631-45.

42. **Elkins, C., C. J. Chen, and C. E. Thomas.** 1995. Characterization of the hgbA locus encoding a hemoglobin receptor from *Haemophilus ducreyi*. Infect Immun **63:**2194-200.
43. **Elkins, C., K. J. Morrow, Jr., and B. Olsen.** 2000. Serum resistance in *Haemophilus ducreyi* requires outer membrane protein DsrA. Infect Immun **68:**1608-19.
44. **Fast, M. V., H. Nsanze, L. J. D'Costa, P. Karasira, I. W. Maclean, P. Piot, W. L. Albritton, and A. R. Ronald.** 1983. Antimicrobial therapy of chancroid: an evaluation of five treatment regimens correlated with in vitro sensitivity. Sex Transm Dis **10:**1-6.
45. **Flashner, Y., E. Mamroud, A. Tidhar, R. Ber, M. Aftalion, D. Gur, S. Lazar, A. Zvi, T. Bino, N. Ariel, B. Velan, A. Shafferman, and S. Cohen.** 2004. Generation of *Yersinia pestis* attenuated strains by signature-tagged mutagenesis in search of novel vaccine candidates. Infect Immun **72:**908-15.
46. **Fortney, K. R., R. S. Young, M. E. Bauer, B. P. Katz, A. F. Hood, R. S. Munson, Jr., and S. M. Spinola.** 2000. Expression of peptidoglycan-associated lipoprotein is required for virulence in the human model of *Haemophilus ducreyi* infection. Infect Immun **68:**6441-8.
47. **Foulongne, V., G. Bourg, C. Cazevieille, S. Michaux-Charachon, and D. O'Callaghan.** 2000. Identification of *Brucella suis* genes affecting intracellular survival in an in vitro human macrophage infection model by signature-tagged transposon mutagenesis. Infect Immun **68:**1297-303.
48. **Frisk, A., C. A. Ison, and T. Lagergard.** 1998. GroEL heat shock protein of *Haemophilus ducreyi*: association with cell surface and capacity to bind to eukaryotic cells. Infect Immun **66:**1252-7.
49. **Garcia Vescovi, E., F. C. Soncini, and E. A. Groisman.** 1996. Mg2+ as an extracellular signal: environmental regulation of *Salmonella* virulence. Cell **84:**165-74.
50. **Ge, Z., Q. Jiang, M. S. Kalisiak, and D. E. Taylor.** 1997. Cloning and functional characterization of *Helicobacter pylori* fumarate reductase operon comprising three structural genes coding for subunits C, A and B. Gene **204:**227-34.
51. **Genevaux, P., C. Georgopoulos, and W. L. Kelley.** 2007. The Hsp70 chaperone machines of *Escherichia coli*: a paradigm for the repartition of chaperone functions. Mol Microbiol **66:**840-57.
52. **Geoffroy, M. C., S. Floquet, A. Metais, X. Nassif, and V. Pelicic.** 2003. Large-scale analysis of the meningococcus genome by gene disruption: resistance to complement-mediated lysis. Genome Res **13:**391-8.
53. **Hansen, E. J., S. R. Lumbley, J. A. Richardson, B. K. Purcell, M. K. Stevens, L. D. Cope, J. Datte, and J. D. Radolf.** 1994. Induction of protective immunity to *Haemophilus ducreyi* in the temperature-dependent rabbit model of experimental chancroid. J Immunol **152:**184-92.
54. **Hayes, F.** 2003. Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. Science **301:**1496-9.

55. **Hensel, M., J. E. Shea, C. Gleeson, M. D. Jones, E. Dalton, and D. W. Holden.** 1995. Simultaneous identification of bacterial virulence genes by negative selection. Science **269:**400-3.
56. **Heungens, K., C. E. Cowles, and H. Goodrich-Blair.** 2002. Identification of *Xenorhabdus nematophila* genes required for mutualistic colonization of *Steinernema carpocapsae* nematodes. Mol Microbiol **45:**1337-53.
57. **Hobbs, M. M., L. R. San Mateo, P. E. Orndorff, G. Almond, and T. H. Kawula.** 1995. Swine model of *Haemophilus ducreyi* infection. Infect Immun **63:**3094-100.
58. **Hunt, T. A., C. Kooi, P. A. Sokol, and M. A. Valvano.** 2004. Identification of *Burkholderia cenocepacia* genes required for bacterial survival in vivo. Infect Immun **72:**4010-22.
59. **Huynen, M. A., T. Dandekar, and P. Bork.** 1999. Variation and evolution of the citric-acid cycle: a genomic perspective. Trends Microbiol **7:**281-91.
60. **Ibba, M., and D. Soll.** 2004. Aminoacyl-tRNAs: setting the limits of the genetic code. Genes Dev **18:**731-8.
61. **Inoue, J., M. Honda, S. Ikawa, T. Shibata, and T. Mikawa.** 2008. The process of displacing the single-stranded DNA-binding protein from single-stranded DNA by RecO and RecR proteins. Nucleic Acids Res **36:**94-109.
62. **Ison, C. A., J. A. Dillon, and J. W. Tapsall.** 1998. The epidemiology of global antibiotic resistance among *Neisseria gonorrhoeae* and *Haemophilus ducreyi.* Lancet **351 Suppl 3:**8-11.
63. **Jacquin-Becker, C., I. Ahel, A. Ambrogelly, B. Ruan, D. Soll, and C. Stathopoulos.** 2002. Cysteinyl-tRNA formation and prolyl-tRNA synthetase. FEBS Lett **514:**34-6.
64. **Janowicz, D., N. R. Luke, K. R. Fortney, B. P. Katz, A. A. Campagnari, and S. M. Spinola.** 2006. Expression of OmpP2A and OmpP2B is not required for pustule formation by *Haemophilus ducreyi* in human volunteers. Microb Pathog **40:**110-5.
65. **Janowicz, D. M., K. R. Fortney, B. P. Katz, J. L. Latimer, K. Deng, E. J. Hansen, and S. M. Spinola.** 2004. Expression of the LspA1 and LspA2 proteins by *Haemophilus ducreyi* is required for virulence in human volunteers. Infect Immun **72:**4528-33.
66. **Johnson, S. R., D. H. Martin, C. Cammarata, and S. A. Morse.** 1995. Alterations in sample preparation increase sensitivity of PCR assay for diagnosis of chancroid. J Clin Microbiol **33:**1036-8.
67. **Joseph, N., V. Duppatla, and D. N. Rao.** 2005. Functional characterization of the DNA mismatch binding protein MutS from *Haemophilus influenzae*. Biochem Biophys Res Commun **334:**891-900.
68. **Jung, H.** 2001. Towards the molecular mechanism of Na(+)/solute symport in prokaryotes. Biochim Biophys Acta **1505:**131-43.
69. **Kaul, R., J. Kimani, N. J. Nagelkerke, F. A. Plummer, J. J. Bwayo, R. C. Brunham, E. N. Ngugi, and A. Ronald.** 1997. Risk factors for genital ulcerations in Kenyan sex workers. The role of human immunodeficiency virus type 1 infection. Sex Transm Dis **24:**387-92.

70. **Kavermann, H., B. P. Burns, K. Angermuller, S. Odenbreit, W. Fischer, K. Melchers, and R. Haas.** 2003. Identification and characterization of *Helicobacter pylori* genes essential for gastric colonization. J Exp Med **197:**813-22.
71. **Keller, K. L., T. L. Overbeck-Carrick, and D. J. Beck.** 2001. Survival and induction of SOS in *Escherichia coli* treated with cisplatin, UV-irradiation, or mitomycin C are dependent on the function of the RecBC and RecFOR pathways of homologous recombination. Mutat Res **486:**21-9.
72. **Knoll, L. J., G. L. Furie, and J. C. Boothroyd.** 2001. Adaptation of signature-tagged mutagenesis for *Toxoplasma gondii*: a negative screening strategy to isolate genes that are essential in restrictive growth conditions. Mol Biochem Parasitol **116:**11-6.
73. **Konkel, M. E., and K. Tilly.** 2000. Temperature-regulated expression of bacterial virulence genes. Microbes Infect **2:**157-66.
74. **Lai, W., C. Y. Chen, S. A. Morse, Y. Htun, H. G. Fehler, H. Liu, and R. C. Ballard.** 2003. Increasing relative prevalence of HSV-2 infection among men with genital ulcers from a mining community in South Africa. Sex Transm Infect **79:**202-7.
75. **Leduc, I., K. E. Banks, K. R. Fortney, K. B. Patterson, S. D. Billings, B. P. Katz, S. M. Spinola, and C. Elkins.** 2008. Evaluation of the repertoire of the TonB-dependent receptors of *Haemophilus ducreyi* for their role in virulence in humans. J Infect Dis **197:**1103-9.
76. **Lee, B. C.** 1991. Iron sources for *Haemophilus ducreyi*. J Med Microbiol **34:**317-22.
77. **Lehoux, D. E., F. Sanschagrin, and R. C. Levesque.** 1999. Defined oligonucleotide tag pools and PCR screening in signature-tagged mutagenesis of essential genes from bacteria. Biotechniques **26:**473-8, 480.
78. **Lewis, D. A., M. K. Stevens, J. L. Latimer, C. K. Ward, K. Deng, R. Blick, S. R. Lumbley, C. A. Ison, and E. J. Hansen.** 2001. Characterization of *Haemophilus ducreyi* cdtA, cdtB, and cdtC mutants in in vitro and in vivo systems. Infect Immun **69:**5626-34.
79. **Magro, C. M., A. N. Crowson, M. Alfa, A. Nath, A. Ronald, J. O. Ndinya-Achola, and J. Nasio.** 1996. A morphological study of penile chancroid lesions in human immunodeficiency virus (HIV)-positive and -negative African men with a hypothesis concerning the role of chancroid in HIV transmission. Hum Pathol **27:**1066-70.
80. **Makarova, K. S., N. V. Grishin, and E. V. Koonin.** 2006. The HicAB cassette, a putative novel, RNA-targeting toxin-antitoxin system in archaea and bacteria. Bioinformatics **22:**2581-4.
81. **Mandrell, R. E., R. McLaughlin, Y. Aba Kwaik, A. Lesse, R. Yamasaki, B. Gibson, S. M. Spinola, and M. A. Apicella.** 1992. Lipooligosaccharides (LOS) of some *Haemophilus* species mimic human glycosphingolipids, and some LOS are sialylated. Infect Immun **60:**1322-8.
82. **Martin, D. H., and R. P. DiCarlo.** 1994. Recent changes in the epidemiology of genital ulcer disease in the United States. The crack cocaine connection. Sex Transm Dis **21:**S76-80.

83. **Mazurkiewicz, P., C. M. Tang, C. Boone, and D. W. Holden.** 2006. Signature-tagged mutagenesis: barcoding mutants for genome-wide screens. Nat Rev Genet **7:**929-39.
84. **Melaugh, W., N. J. Phillips, A. A. Campagnari, R. Karalus, and B. W. Gibson.** 1992. Partial characterization of the major lipooligosaccharide from a strain of *Haemophilus ducreyi*, the causative agent of chancroid, a genital ulcer disease. J Biol Chem **267:**13434-9.
85. **Melaugh, W., N. J. Phillips, A. A. Campagnari, M. V. Tullius, and B. W. Gibson.** 1994. Structure of the major oligosaccharide from the lipooligosaccharide of *Haemophilus ducreyi* strain 35000 and evidence for additional glycoforms. Biochemistry **33:**13070-8.
86. **Menard, R., P. J. Sansonetti, and C. Parsot.** 1993. Nonpolar mutagenesis of the ipa genes defines IpaB, IpaC, and IpaD as effectors of *Shigella flexneri* entry into epithelial cells. J Bacteriol **175:**5899-906.
87. **Mertz, K. J., J. B. Weiss, R. M. Webb, W. C. Levine, J. S. Lewis, K. A. Orle, P. A. Totten, J. Overbaugh, S. A. Morse, M. M. Currier, M. Fishbein, and M. E. St Louis.** 1998. An investigation of genital ulcers in Jackson, Mississippi, with use of a multiplex polymerase chain reaction assay: high prevalence of chancroid and human immunodeficiency virus infection. J Infect Dis **178:**1060-6.
88. **Mhlanga-Mutangadura, T., G. Morlin, A. L. Smith, A. Eisenstark, and M. Golomb.** 1998. Evolution of the major pilus gene cluster of *Haemophilus influenzae*. J Bacteriol **180:**4693-703.
89. **Miyamoto, A., S. Matsuyama, and H. Tokuda.** 2002. Dominant negative mutant of a lipoprotein-specific molecular chaperone, LolA, tightly associates with LolCDE. FEBS Lett **528:**193-6.
90. **Mohammed, T. T., and Y. M. Olumide.** 2008. Chancroid and human immunodeficiency virus infection--a review. Int J Dermatol **47:**1-8.
91. **Morita, Y., K. Kodama, S. Shiota, T. Mine, A. Kataoka, T. Mizushima, and T. Tsuchiya.** 1998. NorM, a putative multidrug efflux protein, of *Vibrio parahaemolyticus* and its homolog in *Escherichia coli*. Antimicrob Agents Chemother **42:**1778-82.
92. **Morse, S. A.** 1989. Chancroid and *Haemophilus ducreyi*. Clin Microbiol Rev **2:**137-57.
93. **Narita, S., K. Tanaka, S. Matsuyama, and H. Tokuda.** 2002. Disruption of lolCDE, encoding an ATP-binding cassette transporter, is lethal for *Escherichia coli* and prevents release of lipoproteins from the inner membrane. J Bacteriol **184:**1417-22.
94. **Nassif, X., D. Puaoi, and M. So.** 1991. Transposition of Tn1545-delta 3 in the pathogenic *Neisseriae*: a genetic tool for mutagenesis. J Bacteriol **173:**2147-54.
95. **Nelson, R. T., J. Hua, B. Pryor, and J. K. Lodge.** 2001. Identification of virulence mutants of the fungal pathogen *Cryptococcus neoformans* using signature-tagged mutagenesis. Genetics **157:**935-47.
96. **Nilsen, A., M. J. Kasubi, S. C. Mohn, D. Mwakagile, N. Langeland, and L. Haarr.** 2007. Herpes simplex virus infection and genital ulcer disease among patients with sexually transmitted infections in Dar es Salaam, Tanzania. Acta Derm Venereol **87:**355-9.

97. **Odumeru, J. A., G. M. Wiseman, and A. R. Ronald.** 1984. Virulence factors of *Haemophilus ducreyi*. Infect Immun **43:**607-11.
98. **Ojha, S., M. Sirois, and J. I. Macinnes.** 2005. Identification of *Actinobacillus suis* genes essential for the colonization of the upper respiratory tract of swine. Infect Immun **73:**7032-9.
99. **Olson, E. R.** 1993. Influence of pH on bacterial gene expression. Mol Microbiol **8:**5-14.
100. **Orlova, M., J. Newlands, A. Das, A. Goldfarb, and S. Borukhov.** 1995. Intrinsic transcript cleavage activity of RNA polymerase. Proc Natl Acad Sci U S A **92:**4596-600.
101. **Pacello, F., P. R. Langford, J. S. Kroll, C. Indiani, G. Smulevich, A. Desideri, G. Rotilio, and A. Battistoni.** 2001. A novel heme protein, the Cu,Zn-superoxide dismutase from *Haemophilus ducreyi*. J Biol Chem **276:**30326-34.
102. **Palmer, K. L., and R. S. Munson, Jr.** 1995. Cloning and characterization of the genes encoding the hemolysin of *Haemophilus ducreyi*. Mol Microbiol **18:**821-30.
103. **Palmer, K. L., C. T. Schnizlein-Bick, A. Orazi, K. John, C. Y. Chen, A. F. Hood, and S. M. Spinola.** 1998. The immune response to *Haemophilus ducreyi* resembles a delayed-type hypersensitivity reaction throughout experimental infection of human subjects. J Infect Dis **178:**1688-97.
104. **Palmer, K. L., A. C. Thornton, K. R. Fortney, A. F. Hood, R. S. Munson, Jr., and S. M. Spinola.** 1998. Evaluation of an isogenic hemolysin-deficient mutant in the human model of *Haemophilus ducreyi* infection. J Infect Dis **178:**191-9.
105. **Parsons, L. M., R. J. Limberger, and M. Shayegani.** 1997. Alterations in levels of DnaK and GroEL result in diminished survival and adherence of stressed *Haemophilus ducreyi*. Infect Immun **65:**2413-9.
106. **Parsons, L. M., A. L. Waring, and M. Shayegani.** 1992. Molecular analysis of the *Haemophilus ducreyi* groE heat shock operon. Infect Immun **60:**4111-8.
107. **Pedersen, K., A. V. Zavialov, M. Y. Pavlov, J. Elf, K. Gerdes, and M. Ehrenberg.** 2003. The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. Cell **112:**131-40.
108. **Perry, R. D.** 1999. Signature-tagged mutagenesis and the hunt for virulence factors: Response. Trends Microbiol **7:**388-389.
109. **Pickett, C. L., and C. A. Whitehouse.** 1999. The cytolethal distending toxin family. Trends Microbiol **7:**292-7.
110. **Plourde, P. J., F. A. Plummer, J. Pepin, E. Agoki, G. Moss, J. Ombette, A. R. Ronald, M. Cheang, L. D'Costa, and J. O. Ndinya-Achola.** 1992. Human immunodeficiency virus type 1 infection in women attending a sexually transmitted diseases clinic in Kenya. J Infect Dis **166:**86-92.
111. **Potvin, E., D. E. Lehoux, I. Kukavica-Ibrulj, K. L. Richard, F. Sanschagrin, G. W. Lau, and R. C. Levesque.** 2003. In vivo functional genomics of *Pseudomonas aeruginosa* for high-throughput screening of new virulence factors and antibacterial targets. Environ Microbiol **5:**1294-308.
112. **Purcell, B. K., J. A. Richardson, J. D. Radolf, and E. J. Hansen.** 1991. A temperature-dependent rabbit model for production of dermal lesions by *Haemophilus ducreyi*. J Infect Dis **164:**359-67.

113. **Purven, M., E. Falsen, and T. Lagergard.** 1995. Cytotoxin production in 100 strains of *Haemophilus ducreyi* from different geographic locations. FEMS Microbiol Lett **129:**221-4.
114. **Read, T. D., S. W. Satola, and M. M. Farley.** 2000. Nucleotide sequence analysis of hypervariable junctions of *Haemophilus influenzae* pilus gene clusters. Infect Immun **68:**6896-902.
115. **Reed, K. C., and D. A. Mann.** 1985. Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acids Res **13:**7207-21.
116. **Risbud, A., K. Chan-Tack, D. Gadkari, R. R. Gangakhedkar, M. E. Shepherd, R. Bollinger, S. Mehendale, C. Gaydos, A. Divekar, A. Rompalo, and T. C. Quinn.** 1999. The etiology of genital ulcer disease by multiplex polymerase chain reaction and relationship to HIV infection among patients attending sexually transmitted disease clinics in Pune, India. Sex Transm Dis **26:**55-62.
117. **Saenz, H. L., and C. Dehio.** 2005. Signature-tagged mutagenesis: technical advances in a negative selection method for virulence gene identification. Curr Opin Microbiol **8:**612-9.
118. **San Mateo, L. R., K. L. Toffer, and T. H. Kawula.** 1998. The sodA gene of *Haemophilus ducreyi* encodes a hydrogen peroxide-inhibitable superoxide dismutase. Gene **207:**251-7.
119. **San Mateo, L. R., K. L. Toffer, P. E. Orndorff, and T. H. Kawula.** 1999. Neutropenia restores virulence to an attenuated Cu,Zn superoxide dismutase-deficient *Haemophilus ducreyi* strain in the swine model of chancroid. Infect Immun **67:**5345-51.
120. **Schonherr, R., M. Hilger, S. Broer, R. Benz, and V. Braun.** 1994. Interaction of *Serratia marcescens* hemolysin (ShlA) with artificial and erythrocyte membranes. Demonstration of the formation of aqueous multistate channels. Eur J Biochem **223:**655-63.
121. **Schulte, J. M., and G. P. Schmid.** 1995. Recommendations for treatment of chancroid, 1993. Clin Infect Dis **20 Suppl 1:**S39-46.
122. **Shah, D. H., M. J. Lee, J. H. Park, J. H. Lee, S. K. Eo, J. T. Kwon, and J. S. Chae.** 2005. Identification of *Salmonella gallinarum* virulence genes in a chicken infection model using PCR-based signature-tagged mutagenesis. Microbiology **151:**3957-68.
123. **Sheehan, B. J., J. T. Bosse, A. J. Beddek, A. N. Rycroft, J. S. Kroll, and P. R. Langford.** 2003. Identification of *Actinobacillus pleuropneumoniae* genes important for survival during infection in its natural host. Infect Immun **71:**3960-70.
124. **Shimoda, Y., H. Mitsui, H. Kamimatsuse, K. Minamisawa, E. Nishiyama, Y. Ohtsubo, Y. Nagata, M. Tsuda, S. Shinpo, A. Watanabe, M. Kohara, M. Yamada, Y. Nakamura, S. Tabata, and S. Sato.** 2008. Construction of signature-tagged mutant library in *Mesorhizobium loti* as a powerful tool for functional genomics. DNA Res **15:**297-308.
125. **Smirnova, A., H. Li, H. Weingart, S. Aufhammer, A. Burse, K. Finis, A. Schenk, and M. S. Ullrich.** 2001. Thermoregulated expression of virulence factors in plant-associated bacteria. Arch Microbiol **176:**393-9.

126. **Smith, J. L., and D. O. Bayles.** 2006. The contribution of cytolethal distending toxin to bacterial pathogenesis. Crit Rev Microbiol **32:**227-48.
127. **Sorg, I., U. M. Goehring, K. Aktories, and G. Schmidt.** 2001. Recombinant *Yersinia* YopT leads to uncoupling of RhoA-effector interaction. Infect Immun **69:**7535-43.
128. **Sousa, M. C., and D. B. McKay.** 2001. Structure of the universal stress protein of *Haemophilus influenzae*. Structure **9:**1135-41.
129. **Spinola, S. M., A. Castellazzo, M. Shero, and M. A. Apicella.** 1990. Characterization of pili expressed by *Haemophilus ducreyi*. Microb Pathog **9:**417-26.
130. **Spinola, S. M., L. M. Wild, M. A. Apicella, A. A. Gaspari, and A. A. Campagnari.** 1994. Experimental human infection with *Haemophilus ducreyi*. J Infect Dis **169:**1146-50.
131. **Steen, R.** 2001. Eradicating chancroid. Bull World Health Organ **79:**818-26.
132. **Stevens, M. K., L. D. Cope, J. D. Radolf, and E. J. Hansen.** 1995. A system for generalized mutagenesis of *Haemophilus ducreyi*. Infect Immun **63:**2976-82.
133. **Stevens, M. K., D. J. Hassett, J. D. Radolf, and E. J. Hansen.** 1996. Cloning and sequencing of the gene encoding the Cu,Zn-superoxide dismutase of *Haemophilus ducreyi*. Gene **183:**35-40.
134. **Stevens, M. K., J. L. Latimer, S. R. Lumbley, C. K. Ward, L. D. Cope, T. Lagergard, and E. J. Hansen.** 1999. Characterization of a *Haemophilus ducreyi* mutant deficient in expression of cytolethal distending toxin. Infect Immun **67:**3900-8.
135. **Stevens, M. K., S. Porcella, J. Klesney-Tait, S. Lumbley, S. E. Thomas, M. V. Norgard, J. D. Radolf, and E. J. Hansen.** 1996. A hemoglobin-binding outer membrane protein is involved in virulence expression by *Haemophilus ducreyi* in an animal model. Infect Immun **64:**1724-35.
136. **Strauss, E. J., and S. Falkow.** 1997. Microbial pathogenesis: genomics and beyond. Science **276:**707-12.
137. **Su, J., J. Yang, D. Zhao, T. H. Kawula, J. A. Banas, and J. R. Zhang.** 2007. Genome-wide identification of *Francisella tularensis* virulence determinants. Infect Immun **75:**3089-101.
138. **Sun, S., N. K. Scheffler, B. W. Gibson, J. Wang, and R. S. Munson, Jr.** 2002. Identification and characterization of the N-acetylglucosamine glycosyltransferase gene of *Haemophilus ducreyi*. Infect Immun **70:**5887-92.
139. **Taniguchi, N., and H. Tokuda.** 2008. Molecular events involved in a single cycle of ligand transfer from an ATP binding cassette transporter, LolCDE, to a molecular chaperone, LolA. J Biol Chem **283:**8538-44.
140. **Thomas, C. E., B. Olsen, and C. Elkins.** 1998. Cloning and characterization of tdhA, a locus encoding a TonB-dependent heme receptor from *Haemophilus ducreyi*. Infect Immun **66:**4254-62.
141. **Thune, R. L., D. H. Fernandez, J. L. Benoit, M. Kelly-Smith, M. L. Rogge, N. J. Booth, C. A. Landry, and R. A. Bologna.** 2007. Signature-tagged mutagenesis of *Edwardsiella ictaluri* identifies virulence-related genes, including a salmonella pathogenicity island 2 class of type III secretion systems. Appl Environ Microbiol **73:**7934-46.

142. **Tian, J., R. Bryk, M. Itoh, M. Suematsu, and C. Nathan.** 2005. Variant tricarboxylic acid cycle in *Mycobacterium tuberculosis*: identification of alpha-ketoglutarate decarboxylase. Proc Natl Acad Sci U S A **102:**10670-5.
143. **Totten, P. A., J. M. Kuypers, C. Y. Chen, M. J. Alfa, L. M. Parsons, S. M. Dutro, S. A. Morse, and N. B. Kiviat.** 2000. Etiology of genital ulcer disease in Dakar, Senegal, and comparison of PCR and serologic assays for detection of *Haemophilus ducreyi*. J Clin Microbiol **38:**268-73.
144. **Totten, P. A., W. R. Morton, G. H. Knitter, A. M. Clark, N. B. Kiviat, and W. E. Stamm.** 1994. A primate model for chancroid. J Infect Dis **169:**1284-90.
145. **Totten, P. A., D. V. Norn, and W. E. Stamm.** 1995. Characterization of the hemolytic activity of *Haemophilus ducreyi*. Infect Immun **63:**4409-16.
146. **Trees, D. L., and S. A. Morse.** 1995. Chancroid and *Haemophilus ducreyi*: an update. Clin Microbiol Rev **8:**357-75.
147. **Unsworth, K. E., and D. W. Holden.** 2000. Identification and analysis of bacterial virulence genes in vivo. Philos Trans R Soc Lond B Biol Sci **355:**613-22.
148. **van Diemen, P. M., F. Dziva, M. P. Stevens, and T. S. Wallis.** 2005. Identification of enterohemorrhagic *Escherichia coli* O26:H- genes required for intestinal colonization in calves. Infect Immun **73:**1735-43.
149. **Ward, C. K., J. L. Latimer, J. Nika, M. Vakevainen, J. R. Mock, K. Deng, R. J. Blick, and E. J. Hansen.** 2003. Mutations in the lspA1 and lspA2 genes of *Haemophilus ducreyi* affect the virulence of this pathogen in an animal model system. Infect Immun **71:**2478-86.
150. **Ward, C. K., J. R. Mock, and E. J. Hansen.** 2004. The LspB protein is involved in the secretion of the LspA1 and LspA2 proteins by *Haemophilus ducreyi*. Infect Immun **72:**1874-84.
151. **Wising, C., L. Molne, I. M. Jonsson, K. Ahlman, and T. Lagergard.** 2005. The cytolethal distending toxin of *Haemophilus ducreyi* aggravates dermal lesions in a rabbit model of chancroid. Microbes Infect **7:**867-74.
152. **Wood, G. E., S. M. Dutro, and P. A. Totten.** 1999. Target cell range of *Haemophilus ducreyi* hemolysin and its involvement in invasion of human epithelial cells. Infect Immun **67:**3740-9.
153. **Xu, X. J., X. Z. Su, Y. Morita, T. Kuroda, T. Mizushima, and T. Tsuchiya.** 2003. Molecular cloning and characterization of the HmrM multidrug efflux pump from *Haemophilus influenzae* Rd. Microbiol Immunol **47:**937-43.
154. **Young, R. S., K. R. Fortney, V. Gelfanova, C. L. Phillips, B. P. Katz, A. F. Hood, J. L. Latimer, R. S. Munson, Jr., E. J. Hansen, and S. M. Spinola.** 2001. Expression of cytolethal distending toxin and hemolysin is not required for pustule formation by *Haemophilus ducreyi* in human volunteers. Infect Immun **69:**1938-42.

APPENDIX: Mutant Sequences and BlastX Searches

Hd12-10-(27) ***aspS***

SEQUENCE:

ACGAACTGAACCACCACCGACTTCAATAGCCGTTGATAACCATATCATAGGCATTAGCAACCGCT
TTCCGATTGGATCGGCTGCTAATTGGTCAGGTGAAAGATCTTTGGGTGCAGTAAATGGATGATGC
ATTGCAGATAAATTGCCTTCTT

BLASTX:

Distance tree of results NEW

```
                                                                    Score
E
Sequences producing significant alignments:                         (Bits)
Value

gb|AE017143.1|  Haemophilus ducreyi strain 35000HP complete ge...   257
8e-66
```

Alignments

```
>  gb|AE017143.1|  Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

 Features in this part of subject sequence:
   aspartyl-tRNA synthetase

 Score =  257 bits (139),  Expect = 8e-66
 Identities = 147/150 (98%), Gaps = 3/150 (2%)
 Strand=Plus/Plus

Query  1
ACGAACTGAACCACCACCGACTTCAATAGCCGTTGATAACCATATCATAGGCATTAGCAA  60
               ||||||||||||||||||||||||
||||||||||||||||||||||||||||||||||||||
Sbjct  459209  ACGAACTGAACCACCACCGACTTC-
ATAGCCGTTGATAACCATATCATAGGCATTAGCAA  459267

Query  61
CCGCTTTCCGATTGGATCGGCTGCTAATTGGTCAGGTGAAAGATCTTTGGGTGCAGTAAA  120
               |||||||
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  459268  CCGCTTT--
GATTGGATCGGCTGCTAATTGGTCAGGTGAAAGATCTTTGGGTGCAGTAAA  459325

Query  121     TGGATGATGCATTGCAGATAAATTGCCTTC  150
               |||||||||||||||||||||||||||||
Sbjct  459326  TGGATGATGCATTGCAGATAAATTGCCTTC  459355
```

Hd1-1(22) ***proS***

SEQUENCE:

TTTTTTCAATTGCTATAATCAAATTGTGCAACCAATTCTGCAATTGTTTTAGCATTTGGCGTATCA
ATCAACACCATCGCTTCAGTTGCTGGTTTAGACTCACCCATCTTA

BLASTX:

Distance tree of results NEW

```
                                                                    Score
E
Sequences producing significant alignments:                         (Bits)
Value

gb|AE017143.1|  Haemophilus ducreyi strain 35000HP complete ge...   193
1e-46
```

Alignments

```
>  gb|AE017143.1|  Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

 Features in this part of subject sequence:
   prolyl-tRNA synthetase

 Score =  193 bits (104),  Expect = 1e-46
 Identities = 107/108 (99%),  Gaps = 1/108 (0%)
 Strand=Plus/Plus

Query  1
TTTTTTCAATTGCTATAATCAAATTGTGCAACCAATTCTGCAATTGTTTTAGCATTTGGC  60
                 ||||||||||||||
||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1613484  TTTTTTCAATTGCTA-
AATCAAATTGTGCAACCAATTCTGCAATTGTTTTAGCATTTGGC  1613542

Query  61       GTATCAATCAACACCATCGCTTCAGTTGCTGGTTTAGACTCACCCATC  108
                 ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1613543  GTATCAATCAACACCATCGCTTCAGTTGCTGGTTTAGACTCACCCATC  1613590
```

Hd9-8(24) *lgtA*

SEQUENCE:

CGGTTATTGCGCAAACCTATAAAAACCTTGAGATTATTGTTGTTAATGATGGTTCAACAGATGGA
ACATTAGCTAAATTACGTCAATTTGAAGCTAAGGATCCAAGAGTCAAGATTATTGATAATATTGT
CAATCAAGGGACGAGTAAATCATTAAATATTGGCATTCAATATTGCCAAGGGGAAATCATTGCAC
GTACAGATTCAGATGATATTGTCGATATACATTGGATNTGAAACATTACATGCGTGAGCTAGATA
ATTCACCGGAGACGATTGCTATTAGTGCTTATTTAGAGTTTTCCTAGCAGAAAAAGGCAATGGTA
GTAAATTAAGTAGATCGCGTAAGCATGGTAAGAACGCCGAAAATCCAATCTCGTCGGAGGCAAT
TATC

BLASTX:

gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
lipooligosaccharide N-acetylglucosamine glycosyltransferase

Score = 697 bits (377), Expect = 0.0
Identities = 387/391 (98%), Gaps = 4/391 (1%)
Strand=Plus/Plus

```
Query  1
CGGTTATTGCGCAAACCTATAAAAACCTTGAGATTATTGTTGTTAATGATGGTTCAACAG  60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  367292
CGGTTATTGCGCAAACCTATAAAAACCTTGAGATTATTGTTGTTAATGATGGTTCAACAG  367351

Query  61
ATGGAACATTAGCTAAATTACGTCAATTTGAAGCTAAGGATCCAAGAGTCAAGATTATTG  120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  367352
ATGGAACATTAGCTAAATTACGTCAATTTGAAGCTAAGGATCCAAGAGTCAAGATTATTG  367411

Query  121
ATAATATTGTCAATCAAGGGACGAGTAAATCATTAAATATTGGCATTCAATATTGCCAAG  180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  367412
ATAATATTGTCAATCAAGGGACGAGTAAATCATTAAATATTGGCATTCAATATTGCCAAG  367471

Query  181
GGGAAATCATTGCACGTACAGATTCAGATGATATTGTCGATATACATTGGATNTGAAACA  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
|||||||
Sbjct  367472  GGGAAATCATTGCACGTACAGATTCAGATGATATTGTCGATATACATTGGAT-
TGAAACA  367530
```

```
Query  241
TTACATGCGTGAGCTAGATAATTCACCGGAGACGATTGCTATTAGTGCTTATTTAGAGTT  300
                  |||
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  367531  TTA-
ATGCGTGAGCTAGATAATTCACCGGAGACGATTGCTATTAGTGCTTATTTAGAGTT  367589

Query  301
TTCCTAGCAGAAAAAGGCAATGGTAGTAAATTAAGTAGATCGCGTAAGCATGGTAAGAAC  360
                  ||
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  367590  TT--
TAGCAGAAAAAGGCAATGGTAGTAAATTAAGTAGATCGCGTAAGCATGGTAAGAAC  367647

Query  361     GCCGAAAATCCAATCTCGTCGGAGGCAATTA  391
                ||||||||||||||||||||||||||||||
Sbjct  367648  GCCGAAAATCCAATCTCGTCGGAGGCAATTA  367678
```

Hd2-2(18) *ccmH*

SEQUENCE:

ATGAAAAAACACTTAGCTATTTTGTTATTAACTGCTTCAGTAGCAACGGCAGCACCGGTTGAATT
TAATCAATTCGATAACGCACAACAGGAAGCGGATTATCAGGCGTTAATTCAAGAATTACGTTGTC
CTCAATGTCAAAATAATAATATTGCTGATTCAAATGCAACAATTGCAACAGATATGCGGACATAA
AGACGTTGGAATTGTTAAAATCAAGGAAAAACTAAACCTGAAGTGGTGGATTATATGATTGAGC
CC

BLASTX:

gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features flanking this part of subject sequence:
2 bp at 5' side: probable thiol:disulfide interchange protein DsbE
1123 bp at 3' side: molybdopterin biosynthesis protein

Score = 459 bits (248), Expect = 4e-126
Identities = 256/259 (98%), Gaps = 3/259 (1%)
Strand=Plus/Plus

```
Query  1      ATGAAAAAACACTTAGCTATTTTGTTATTAACTGCTTCAGTAGCAACGGCAGCACCGGTT
60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  51434  ATGAAAAAACACTTAGCTATTTTGTTATTAACTGCTTCAGTAGCAACGGCAGCACCGGTT
51493

Query  61     GAATTTAATCAATTCGATAACGCACAACAGGAAGCGGATTATCAGGCGTTAATTCAAGAA
120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  51494  GAATTTAATCAATTCGATAACGCACAACAGGAAGCGGATTATCAGGCGTTAATTCAAGAA
51553

Query  121    TTACGTTGTCCTCAATGTCAAAATAATAATATTGCTGATTCAAATGCAACAATTGCAACA
180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  51554  TTACGTTGTCCTCAATGTCAAAATAATAATATTGCTGATTCAAATGCAACAATTGCAACA
51613

Query  181    GATATGCGGACATAAAGACGTTGGAATTGTTAAAATCAAGGAAAAACTAAACCTGAAGTG
240
              |||||||| |||||| |||||||||||||||||| ||||||||||||||||||||||||
Sbjct  51614  GATATGCG-ACATAA-GACGTTGGAATTGTTAAAA-CAAGGAAAAACTAAACCTGAAGTG
51670

Query  241    GTGGATTATATGATTGAGC  259
              ||||||||||||||||||
Sbjct  51671  GTGGATTATATGATTGAGC  51689
```

ALIGNMENT BLASTP:

Features in this part of subject sequence:
Cytochrome c-type biogenesis protein ccmH precursor

Score = 75.0 bits (158), Expect = 1e-11
Identities = 33/77 (42%), Positives = 46/77 (59%), Gaps = 0/77 (0%)
Frame = +1/-3

```
Query  25
LLTASVATAAPVEFNQFDNAQQEADYQALIQELRCPQCQNNNIADSNATIATDMRHKTLE  204
               L+ + VA A  V+  QF N        AL + LRCPQCQN N+ +SNAT+A  +R +
E
Sbjct  1253755
LVFSLVAKAEMVDTFQFQNETDRVRAVALAKSLRCPQCQNQNLVESNATMAYKLRLEVYE  1253576

Query  205      LLKQGKTKPEVVDYMIE  255
                ++ QGKT  E++  M E
Sbjct  1253575  MVNQGKTDEEIIKIMTE  1253525
```

Hd3-5(26) *recR*

SEQUENCE:

ATTCTTGTTCTAGCCGTTGTTGTAATAAATCTAAGCCGATTTCACGTGGACCAATGCCATCTAATG
GCGATAAATGGCCCATTAATACAAAATAACGGCCAGAAAATTGACCGGTTTGTTCGATTGCTTGG
ATATCTTCTGGCATTTCTACTATGCAAAGTTGACCGCTTATTTGGCGACGTGGATTTTTACAAATT
GTACATTCTTCTTCTTCGGTGAATGTACGACAAGCATTGCAATGACCAATATGCGTCATTGCTTCA
TTTAATGCTTTGGCTAGATTCAAGCCACCATTTCTAATTTCGTTGTAATAAATGATAAGCCATACG
TTGTGCAGATTTTGGGCCTATGGCCAGGTAAAGCACGTAACGCCAAT

BLASTX:

gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
recombination DNA repair protein

Score = 675 bits (365), Expect = 0.0
Identities = 370/372 (99%), Gaps = 2/372 (0%)
Strand=Plus/Plus

```
Query  1
ATTCTTGTTCTAGCCGTTGTTGTAATAAATCTAAGCCGATTTCACGTGGACCAATGCCAT  60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  258735
ATTCTTGTTCTAGCCGTTGTTGTAATAAATCTAAGCCGATTTCACGTGGACCAATGCCAT  258794

Query  61
CTAATGGCGATAAATGGCCCATTAATACAAAATAACGGCCAGAAAATTGACCGGTTTGTT  120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  258795
CTAATGGCGATAAATGGCCCATTAATACAAAATAACGGCCAGAAAATTGACCGGTTTGTT  258854

Query  121
CGATTGCTTGGATATCTTCTGGCATTTCTACTATGCAAAGTTGACCGCTTATTTGGCGAC  180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  258855
CGATTGCTTGGATATCTTCTGGCATTTCTACTATGCAAAGTTGACCGCTTATTTGGCGAC  258914

Query  181
GTGGATTTTTACAAATTGTACATTCTTCTTCTTCGGTGAATGTACGACAAGCATTGCAAT  240

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  258915
GTGGATTTTTACAAATTGTACATTCTTCTTCTTCGGTGAATGTACGACAAGCATTGCAAT  258974

Query  241
GACCAATATGCGTCATTGCTTCATTTAATGCTTTGGCTAGATTCAAGCCACCATTTCTAA  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  258975
GACCAATATGCGTCATTGCTTCATTTAATGCTTTGGCTAGATTCAAGCCACCATTTCTA-  259033

Query  301
TTTCGTTGTAATAAATGATAAGCCATACGTTGTGCAGATTTTGGGCCTATGGCCAGGTAA  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
|||||||
Sbjct  259034  TTTCGTTGTAATAAATGATAAGCCATACGTTGTGCAGATTTTGGGCCTATG-
CCAGGTAA  259092

Query  361     AGCACGTAACGC  372
               ||||||||||||
Sbjct  259093  AGCACGTAACGC  259104
```

Hd6-12(16)*frdA*

SEQUENCE:

CTGCGGTGATTAAAGATAGCGATTCATTTGATAATCATTTTAATGATACTGTCGGGGGTGGGGAC
TGGCTGTGTGAACAAGATATTGTGGAATATTTTGTTGAACATTCACCGCTTGAAATGACTCAATTA
GAACGTTGGGGTTGTCCTTGGAGCCGTCGAGCAGATGGCGAAGTTAATGTGCGTCGTTTTGGTGG
GATGAAAATTGAACGTACTTGGTTCGCTGCAGATAAAACAGGCTTTCACTTATTACATACCCTTTT
CCAAACATCGATTAAATACCCTAACATTATTCGGTTTGATGAGCATTTTGTGTTAGATATTTTGAC
CGATAATGGTGAAGCGCGAGGCTGTGTTGCGATGAATATGTTAAGC

BLASTX:

gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
fumarate reductase flavoprotein subunit

Score = 680 bits (368), Expect = 0.0
Identities = 368/368 (100%), Gaps = 0/368 (0%)
Strand=Plus/Plus

```
Query  1      CTGCGGTGATTAAAGATAGCGATTCATTTGATAATCATTTTAATGATACTGTCGGGGGTG
60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  28744  CTGCGGTGATTAAAGATAGCGATTCATTTGATAATCATTTTAATGATACTGTCGGGGGTG
28803

Query  61     GGGACTGGCTGTGTGAACAAGATATTGTGGAATATTTTGTTGAACATTCACCGCTTGAAA
120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  28804  GGGACTGGCTGTGTGAACAAGATATTGTGGAATATTTTGTTGAACATTCACCGCTTGAAA
28863

Query  121    TGACTCAATTAGAACGTTGGGGTTGTCCTTGGAGCCGTCGAGCAGATGGCGAAGTTAATG
180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  28864  TGACTCAATTAGAACGTTGGGGTTGTCCTTGGAGCCGTCGAGCAGATGGCGAAGTTAATG
28923

Query  181    TGCGTCGTTTTGGTGGGATGAAAATTGAACGTACTTGGTTCGCTGCAGATAAAACAGGCT
240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  28924  TGCGTCGTTTTGGTGGGATGAAAATTGAACGTACTTGGTTCGCTGCAGATAAAACAGGCT
28983

Query  241    TTCACTTATTACATACCCTTTTCCAAACATCGATTAAATACCCTAACATTATTCGGTTTG
300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  28984  TTCACTTATTACATACCCTTTTCCAAACATCGATTAAATACCCTAACATTATTCGGTTTG
29043
```

```
Query  301    ATGAGCATTTTGTGTTAGATATTTTGACCGATAATGGTGAAGCGCGAGGCTGTGTTGCGA
360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  29044  ATGAGCATTTTGTGTTAGATATTTTGACCGATAATGGTGAAGCGCGAGGCTGTGTTGCGA
29103

Query  361    TGAATATG  368
              ||||||||
Sbjct  29104  TGAATATG  29111
```

Hd7-2(01) *sucA*

SEQUENCE:

ATTCTTCCCCATTTTTGTTCGCCTGAACTAATAAATTGGTCAATCACAATTTGAGCACCATTCGCA
AAGTCACCAAATTGTGCTTCCCAGATGGTTAAGGTTTTAGGATCTGTTATAGCATAGCCATACTCA
AATGCTAATACAGCTTCTTCAGAGAGGACGGAATCCCAGACTTCAAAGCGAGCTTGTTGAGCATG
TAAATGTGTTAAAGGTATATAATTGCTTCCATCTTGTTGATTATGAATAACGGCATGCCGATGGA
AAAATGTACCGCGGCCGGCATCTTCGCCAGACAACCGAATCGGGATACCTTCATCTAATAAGGTG
GCATACGCCATCGTTTCTGCCATTCCCCAATCAAATAATTTTTGGCCTTGATACATTTGTTTGCGG
TCGTGGTAGATTTTTTCAACCCTTGAATGTAGCTGTAAATGCTCAGGAAGCTGACAAAGTCGTTC
GGCTAAAGTGTTAAAGCGTTCACTAGCAAATTTGCTCTGATACGGAGCGCTTAAATCATCATTGA
GATATTGTAACCACTCGCTTTTCGTCTTGTCGAGCGCGCGG

BLASTX:

gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
2-oxoglutarate dehydrogenase E1 component

Score = 1042 bits (564), Expect = 0.0
Identities = 564/564 (100%), Gaps = 0/564 (0%)
Strand=Plus/Plus

```
Query  1
ATTCTTCCCCATTTTTGTTCGCCTGAACTAATAAATTGGTCAATCACAATTTGAGCACCA  60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1092855
ATTCTTCCCCATTTTTGTTCGCCTGAACTAATAAATTGGTCAATCACAATTTGAGCACCA  1092914

Query  61
TTCGCAAAGTCACCAAATTGTGCTTCCCAGATGGTTAAGGTTTTAGGATCTGTTATAGCA  120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1092915
TTCGCAAAGTCACCAAATTGTGCTTCCCAGATGGTTAAGGTTTTAGGATCTGTTATAGCA  1092974

Query  121
TAGCCATACTCAAATGCTAATACAGCTTCTTCAGAGAGGACGGAATCCCAGACTTCAAAG  180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1092975
TAGCCATACTCAAATGCTAATACAGCTTCTTCAGAGAGGACGGAATCCCAGACTTCAAAG  1093034

Query  181
CGAGCTTGTTGAGCATGTAAATGTGTTAAAGGTATATAATTGCTTCCATCTTGTTGATTA  240

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1093035
CGAGCTTGTTGAGCATGTAAATGTGTTAAAGGTATATAATTGCTTCCATCTTGTTGATTA  1093094
```

```
Query  241
TGAATAACGGCATGCCGATGGAAAAATGTACCGCGGCCGGCATCTTCGCCAGACAACCGA  300

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1093095
TGAATAACGGCATGCCGATGGAAAAATGTACCGCGGCCGGCATCTTCGCCAGACAACCGA  1093154

Query  301
ATCGGGATACCTTCATCTAATAAGGTGGCATACGCCATCGTTTCTGCCATTCCCCAATCA  360

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1093155
ATCGGGATACCTTCATCTAATAAGGTGGCATACGCCATCGTTTCTGCCATTCCCCAATCA  1093214

Query  361
AATAATTTTTGGCCTTGATACATTTGTTTGCGGTCGTGGTAGATTTTTTCAACCCTTGAA  420

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1093215
AATAATTTTTGGCCTTGATACATTTGTTTGCGGTCGTGGTAGATTTTTTCAACCCTTGAA  1093274

Query  421
TGTAGCTGTAAATGCTCAGGAAGCTGACAAAGTCGTTCGGCTAAAGTGTTAAAGCGTTCA  480

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1093275
TGTAGCTGTAAATGCTCAGGAAGCTGACAAAGTCGTTCGGCTAAAGTGTTAAAGCGTTCA  1093334

Query  481
CTAGCAAATTTGCTCTGATACGGAGCGCTTAAATCATCATTGAGATATTGTAACCACTCG  540

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1093335
CTAGCAAATTTGCTCTGATACGGAGCGGTTAAATCATCATTGAGATATTGTAACCACTCG  1093394

Query  541       CTTTTCGTCTTGTCGAGCGCGCGG  564
                  ||||||||||||||||||||||||
Sbjct  1093395   CTTTTCGTCTTGTCGAGCGCGCGG  1093418
```

Hd1-12(26) *lpxB*

SEQUENCE:

TAAATTTGTGGCACGTTTAATTTTATGTATGCGATCCTGACGCCATGCCCATACTGAAGGGCTCAC
ATAGTGAATAGTCGTGATGCCCTTTGCTTTCAATTTTTGTTCAATTGTTAAATTGAAATCAGGCGC
GTCAATGCCAATAAATATATCAGGTTGTTGCTGAAGCATTGTTTGAATCACTTGTTTACGGCGTTT
GAGTAAACGCGGTAAATGTTTGACCACTTCAGCCAACCCCATCACCGCCAGCTCTTCCATATCAA
ACAATGTTTGACANACCCGCTTGAATCATTTGAGGACCGGNNCTACGCCAATAAAACGTGCATTT
GGGTAATGAATTTTTAGCGCGTTNGATCAGCCCAGCGCCTAAAATATCACCAGAAATTTCACCAG
CAACAAGCGCGATAAGCGNNNNN

BLASTX:

gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
lipid-A-disaccharide synthase

Score = 734 bits (397), Expect = 0.0
Identities = 407/411 (99%), Gaps = 4/411 (0%)
Strand=Plus/Plus

```
Query  1
TAAATTTGTGGCACGTTTAATTTTATGTATGCGATCCTGACGCCATGCCCATACTGAAGG  60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  673676
TAAATTTGTGGCACGTTTAATTTTATGTATGCGATCCTGACGCCATGCCCATACTGAAGG  673735

Query  61
GCTCACATAGTGAATAGTCGTGATGCCCTTTGCTTTCAATTTTTGTTCAATTGTTAAATT  120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  673736
GCTCACATAGTGAATAGTCGTGATGCCCTTTGCTTTCAATTTTTGTTCAATTGTTAAATT  673795

Query  121
GAAATCAGGCGCGTCAATGCCAATAAATATATCAGGTTGTTGCTGAAGCATTGTTTGAAT  180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  673796
GAAATCAGGCGCGTCAATGCCAATAAATATATCAGGTTGTTGCTGAAGCATTGTTTGAAT  673855

Query  181
CACTTGTTTACGGCGTTTGAGTAAACGCGGTAAATGTTTGACCACTTCAGCCAACCCCAT  240

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  673856
CACTTGTTTACGGCGTTTGAGTAAACGCGGTAAATGTTTGACCACTTCAGCCAACCCCAT  673915
```

```
Query  241
CACCGCCAGCTCTTCCATATCAAACAATGTTTGACANACCCGCTTGAATCATTTGAGGAC  300
                ||||||||||||||||||||||||||||||||||||||||||||||
||||||||||||||||||||||
Sbjct  673916  CACCGCCAGCTCTTCCATATCAAACAATGTTTGACA-
ACCCGCTTGAATCATTTGAGGAC  673974

Query  301
CGGNNCTACGCCAATAAAACGTGCATTTGGGTAATGAATTTTTAGCGCGTTNGATCAGCC  360
                |||  ||||||||||||||||||||||||||||||||||||||||||||||||||
||||||||
Sbjct  673975  CGG--CTACGCCAATAAAACGTGCATTTGGGTAATGAATTTTTAGCGCGTT-
GATCAGCC  674031

Query  361     CAGCGCCTAAAATATCACCAGAAATTTCACCAGCAACAAGCGCGATAAGCG  411
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  674032  CAGCGCCTAAAATATCACCAGAAATTTCACCAGCAACAAGCGCGATAAGCG  674082
```

Hd9-30(20) *lolA*

SEQUENCE:

GTATCCAAAAAACAATCTTGACCGTGCTCTTTTCTTCATTGTCTTCAATTGCCTTTGCCGATATGC
AATCGATAGCAGAATTACAACGTCGTTTAGAGCACGTTGCTCAATATAGTGCGGATTTTGAACAA
ACTGTGCGTTCAAGCAAAGGCCAACAAATTCAATCAGGTCGCGGTAAATTTCAGGTTAAACGTCC
AAACTTGTTTAGAATGGNATATTAACGCACCGCAAGNAAAACGTAATTGTATCNNGGATGGCGA
AAACTTATGGTTTTATGATCCTTTTGTGGCACAAGTTACCGTTAATAGNTGTGCAAAATGCAGTAA
ATGGTATTNN

BLASTX:

> gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
outer membrane lipoprotein carrier protein

Score = 580 bits (314), Expect = 1e-162
Identities = 327/332 (98%), Gaps = 5/332 (1%)
Strand=Plus/Plus

```
Query  1
GTATCCAAAAAACAATCTTGACCGTGCTCTTTTCTTCATTGTCTTCAATTGCCTTTGCCG  60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1004491
GTATCCAAAAAACAATCTTGACCGTGCTCTTTTCTTCATTGTCTTCAATTGCCTTTGCCG  1004550

Query  61
ATATGCAATCGATAGCAGAATTACAACGTCGTTTAGAGCACGTTGCTCAATATAGTGCGG  120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1004551
ATATGCAATCGATAGCAGAATTACAACGTCGTTTAGAGCACGTTGCTCAATATAGTGCGG  1004610

Query  121
ATTTTGAACAAACTGTGCGTTCAAGCAAAGGCCAACAAATTCAATCAGGTCGCGGTAAAT  180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1004611
ATTTTGAACAAACTGTGCGTTCAAGCAAAGGCCAACAAATTCAATCAGGTCGCGGTAAAT  1004670

Query  181
TTCAGGTTAAACGTCCAAACTTGTTTAGAATGGNATATTAACGCACCGCAAGNAAAACGT  240
                ||||||||||||||||||||||||||||||| ||||||||||||||||||
|||||||
Sbjct  1004671  TTCAGGTTAAACGTCCAAACTTGTTTAGAATGG-ATATTAACGCACCGCAAG-
AAAACGT  1004728
```

```
Query  241
AATTGTATCNNGGATGGCGAAAACTTATGGTTTTATGATCCTTTTGTGGCACAAGTTACC  300
                  ||||||||
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1004729  AATTGTATC--
GGATGGCGAAAACTTATGGTTTTATGATCCTTTTGTGGCACAAGTTACC  1004786

Query  301      GTTAATAGNTGTGCAAAATGCAGTAAATGGTA  332
                 |||||||| ||||||||||||||||||||||
Sbjct  1004787  GTTAATAG-TGTGCAAAATGCAGTAAATGGTA  1004817
```

Hd1-2(06) *putP*

SEQUENCE:

**AGCCAATGGCTAACCACCCTTCAATTAACCCACCCGCATAGACCGCACCGGGTAACCCCAT
TAATAGCCAGCCAGACATATCGGATGCGCCAGCGGATAATGCCGTGACGAAACTACCTAAA
CGACGACCGCCTAAAATGTAGTCTGATAAATTATTTGTTACACGGTAAGCGATAAGACCGAT
CAATATCATTCCAATAATGTAGATTAGGAAAGTCGTTGTTGTAGGATCAAATCCAAACAT**

BLASTX:

```
gb|AE017143.1|  Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

 Features in this part of subject sequence:
   sodium/proline symporter, proline permease

 Score =  451 bits (244),  Expect = 6e-124
 Identities = 244/244 (100%), Gaps = 0/244 (0%)
 Strand=Plus/Plus

Query  1
AGCCAATGGCTAACCACCCTTCAATTAACCCACCCGCATAGACCGCACCGGGTAACCCCA  60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  795017
AGCCAATGGCTAACCACCCTTCAATTAACCCACCCGCATAGACCGCACCGGGTAACCCCA  795076

Query  61
TTAATAGCCAGCCAGACATATCGGATGCGCCAGCGGATAATGCCGTGACGAAACTACCTA  120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  795077
TTAATAGCCAGCCAGACATATCGGATGCGCCAGCGGATAATGCCGTGACGAAACTACCTA  795136

Query  121
AACGACGACCGCCTAAAATGTAGTCTGATAAATTATTTGTTACACGGTAAGCGATAAGAC  180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  795137
AACGACGACCGCCTAAAATGTAGTCTGATAAATTATTTGTTACACGGTAAGCGATAAGAC  795196

Query  181
CGATCAATATCATTCCAATAATGTAGATTAGGAAAGTCGTTGTTGTAGGATCAAATCCAA  240

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  795197
CGATCAATATCATTCCAATAATGTAGATTAGGAAAGTCGTTGTTGTAGGATCAAATCCAA  795256

Query  241      ACAT  244
                 ||||
Sbjct  795257   ACAT  795260
```

Hd12-3(20) ***lspA1***

SEQUENCE:

TNCTNAATTTTTTCACTGCTTCTATATGAGGATCAATATTTTCAGCCTTTAATTTAGTTTGTTTAAA
AGCTGCTTCCACATTCGGTAATTTTAGCGGTTTATAGTCAGTATTNGGTAAGCGCTACTTCTACGC
CATTCTCAACTATATGATTTACTCCTTTTGTCTTATCGTTAATTATATCGATATTAGTATTACGATA
ACCTACACTTGTTTGCGCTGTATCCTTTTCAGAACGCTTATTAGAAGAAGATTTATCAAATNAAAT
TTCTTACTCTCTGCAATAAGCTCGGACCTTCTTGTCTAGCTCTACGTGGAGAACTCTTTAAGTCAG
ACTCTGATATAGAAGTCTTCAATGTTTTTAACGGATTCGCAATTACATGGTCTTTCAGCGTAAATT
TTAAATTAGGATTATCTTTATAATTCTCATAGAGATAGCGTAATTCTCTCGATGTCACTGAATCAT
AATGTTCATGATTAGGATTAAAAACATCTTTCAATTTATCCATATTAATCTGCGTATCAGCTCTAA
AACCTGGGCTATTTGTCACNTAAACCATNNCAATGCAAAGTAAAGCTGTCCTCCTCGTTCTAGCG
TTTCTTTCGCGATAAATTTCGTCTTCTCGGTNNN

BLASTX:

> gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
large supernatant protein 1

Score = 1116 bits (604), Expect = 0.0
Identities = 617/622 (99%), Gaps = 5/622 (0%)
Strand=Plus/Plus

```
Query  6
AATTTTTTCACTGCTTCTATATGAGGATCAATATTTTCAGCCTTTAATTTAGTTTGTTTA  65

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1238814
AATTTTTTCACTGCTTCTATATGAGGATCAATATTTTCAGCCTTTAATTTAGTTTGTTTA  1238873

Query  66
AAAGCTGCTTCCACATTCGGTAATTTTAGCGGTTTATAGTCAGTATTNGGTAAGCGCTAC  125
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||
||||||||||||
Sbjct  1238874  AAAGCTGCTTCCACATTCGGTAATTTTAGCGGTTTATAGTCAGTATT-
GGTAAGCGCTAC  1238932

Query  126
TTCTACGCCATTCTCAACTATATGATTTACTCCTTTTGTCTTATCGTTAATTATATCGAT  185

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1238933
TTCTACGCCATTCTCAACTATATGATTTACTCCTTTTGTCTTATCGTTAATTATATCGAT  1238992

Query  186
ATTAGTATTACGATAACCTACACTTGTTTGCGCTGTATCCTTTTCAGAACGCTTATTAGA  245

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1238993
ATTAGTATTACGATAACCTACACTTGTTTGCGCTGTATCCTTTTCAGAACGCTTATTAGA  1239052
```

```
Query  246
AGAAGATTTATCAAATNAAATTTCTTACTCTCTGCAATAAGCTCGGACCTTCTTGTCTAG  305
                  ||||||||||||||||
||||||||||||||||||||||||||||||||||||||||||
Sbjct  1239053  AGAAGATTTATCAAAT-
AAATTTCTTACTCTCTGCAATAAGCTCGGACCTTCTTGTCTAG  1239111

Query  306
CTCTACGTGGAGAACTCTTTAAGTCAGACTCTGATATAGAAGTCTTCAATGTTTTTAACG  365

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1239112
CTCTACGTGGAGAACTCTTTAAGTCAGACTCTGATATAGAAGTCTTCAATGTTTTTAACG  1239171

Query  366
GATTCGCAATTACATGGTCTTTCAGCGTAAATTTTAAATTAGGATTATCTTTATAATTCT  425

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1239172
GATTCGCAATTACATGGTCTTTCAGCGTAAATTTTAAATTAGGATTATCTTTATAATTCT  1239231

Query  426
CATAGAGATAGCGTAATTCTCTCGATGTCACTGAATCATAATGTTCATGATTAGGATTAA  485

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1239232
CATAGAGATAGCGTAATTCTCTCGATGTCACTGAATCATAATGTTCATGATTAGGATTAA  1239291

Query  486
AAACATCTTTCAATTTATCCATATTAATCTGCGTATCAGCTCTAAAACCTGGGCTATTTG  545

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1239292
AAACATCTTTCAATTTATCCATATTAATCTGCGTATCAGCTCTAAAACCTGGGCTATTTG  1239351

Query  546
TCACNTAAACCATNNCTAATGCAAAGTAAAGCTGTCCTCCTCGTTCTAGCGTTTCTTTCG  605
                  ||||  ||||||||
||||||||||||||||||||||||||||||||||||||||||
Sbjct  1239352  TCAC-TAAACCAT--
CTAATGCAAAGTAAAGCTGTCCTCCTCGTTCTAGCGTTTCTTTCG  1239408

Query  606       CGATAAATTTCGTCTTCTCGGT  627
                  ||||||||||||||||||||||
Sbjct  1239409  CGATAAATTTCGTCTTCTCGGT  1239430
```

Hd7-8(06) HD1268

SEQUENCE:

CCGACAACAATTACCTCAGTAATATTTAAATAAGTAAGCAGATCAATCAGGTCTTGTGCCATTAG
GGAATAATTCATCTGTTCAGAATGAAAAGATCGACCATGATTACGCAAATCTACCCGTAGGATAT
TAAATTTTTCCGCAAATTTACGAGCAATAATGCCTAAATTATTCATATCGCCAAATAAACCATGTA
AGAATAACATAGTTTGAGCATTTGGCTGTTGAATTGCAGGCT

BLASTX:

```
>  gb|AE017143.1|  Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

 Features in this part of subject sequence:
   conserved putative esterase/lipase

 Score =  440 bits (238),  Expect = 1e-120
 Identities = 238/238 (100%), Gaps = 0/238 (0%)
 Strand=Plus/Plus

Query  1
CCGACAACAATTACCTCAGTAATATTTAAATAAGTAAGCAGATCAATCAGGTCTTGTGCC  60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1030467
CCGACAACAATTACCTCAGTAATATTTAAATAAGTAAGCAGATCAATCAGGTCTTGTGCC  1030526

Query  61
ATTAGGGAATAATTCATCTGTTCAGAATGAAAAGATCGACCATGATTACGCAAATCTACC  120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1030527
ATTAGGGAATAATTCATCTGTTCAGAATGAAAAGATCGACCATGATTACGCAAATCTACC  1030586

Query  121
CGTAGGATATTAAATTTTTCCGCAAATTTACGAGCAATAATGCCTAAATTATTCATATCG  180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1030587
CGTAGGATATTAAATTTTTCCGCAAATTTACGAGCAATAATGCCTAAATTATTCATATCG  1030646

Query  181      CCAAATAAACCATGTAAGAATAACATAGTTTGAGCATTTGGCTGTTGAATTGCAGGCT
238

                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1030647  CCAAATAAACCATGTAAGAATAACATAGTTTGAGCATTTGGCTGTTGAATTGCAGGCT
1030704
```

Hd1-2(22) HD0464

SEQUENCE:

CTTTTTTTACGCAGTATGCAGAATGGAGTTATTTAGACTCGGCAGCGACCACCTTAAAACCGGCA
GTATTAATAGAGGCTATGAGTGATTTTTATGCTTCAGCGGGATCGGTTCATCGTAGTCAATATGAT
TTAGCCCAAACGCAAGACTATGAGTTAGCGCGGNNCAGTTAGTGGTAAAGTGGTTTAATGTTGAG
TCCGCANGAGGGTGTCATTTGGACTAGTGGCACAACACATCTCTACTAATTCTAGTGGCAAGCGG
ATTAGCTTATAGTTTAAATGCGTGTGATGAAATTATTATTTCGATTGCAGNNN

BLASTX:

Distance tree of results NEW

Score
E
Sequences producing significant alignments: (Bits)
Value

gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete ge... 542
4e-151

Alignments

>gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
putative aminotransferase

Score = 542 bits (293), Expect = 4e-151
Identities = 306/311 (98%), Gaps = 5/311 (1%)
Strand=Plus/Plus

```
Query  1
CttttttACGCAGTATGCAGAATGGAGTTATTTAGACTCGGCAGCGACCACCTTAAAAC  60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  364860
CTTTTTTTACGCAGTATGCAGAATGGAGTTATTTAGACTCGGCAGCGACCACCTTAAAAC  364919

Query  61
CGGCAGTATTAATAGAGGCTATGAGTGATTTTTATGCTTCAGCGGGATCGGTTCATCGTA  120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  364920
CGGCAGTATTAATAGAGGCTATGAGTGATTTTTATGCTTCAGCGGGATCGGTTCATCGTA  364979

Query  121
GTCAATATGATTTAGCCCAAACGCAAGACTATGAGTTAGCGCGGNNCAGTTAGTGGTAAA  180
```

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||
|||||||||||||||
Sbjct  364980  GTCAATATGATTTAGCCCAAACGCAAGACTATGAGTTAGCGCGG--
CAGTTAGTGGTAAA  365037

Query  181
GTGGTTTAATGTTGAGTCCGCANGAGGGTGTCATTTGGACTAGTGGCACAACACATCTCT  240
                |||||||||||||||||||||| ||||||||||||||||||||||||||||||||
|||
Sbjct  365038  GTGGTTTAATGTTGAGTCCGCA-GAGGGTGTCATTTGGACTAGTGGCACAACACAT-
TCT  365095

Query  241
ACTAATTCTAGTGGCAAGCGGATTAGCTTATAGTTTAAATGCGTGTGATGAAATTATTAT  300
                |||||||
|||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  365096  ACTAATT-
TAGTGGCAAGCGGATTAGCTTATAGTTTAAATGCGTGTGATGAAATTATTAT  365154

Query  301      TTCGATTGCAG  311
                ||||||||||||
Sbjct  365155  TTCGATTGCAG  365165
```

Hd10-2(20) *hicB*

SEQUENCE:

CTAACTAATAGTCACTTTTCCTGCANNGAATATTGCACCTCATCTTCAAATAAAGAAAAGCTNAA
TAGATTTATTTACTTNGAGTATGAATATTAGATTCAACGATAATGCCATAGTAGTTGAAATTGAAT
TCCTTGCANTGTATTGATAAACGGTAGATACCTCTCTCTCATAAGAGAGCACTTCATTTTCAATGG
CCTTATAGACAAAAGCGTTAAGTGCTAATCCTTGCNN

BLASTX:

Distance tree of results NEW

Score
E
Sequences producing significant alignments: (Bits)
Value

gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete ge... 370
2e-99

Alignments

> gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
HicB protein

Score = 370 bits (200), Expect = 2e-99
Identities = 223/232 (96%), Gaps = 9/232 (3%)
Strand=Plus/Plus

```
Query  1
CTAACTAATAGTCACTTTTCCTGCANNGAATATTGCACCTCATCTTCAAATAAAGAAAAG  60
               ||||||||||||||||||||||||
||||||||||||||||||||||||||||||||
Sbjct  728678  CTAACTAATAGTCACTTTTCCTGCA--
GAATATTGCACCTCATCTTCAAATAAAGAAAAG  728735

Query  61
CTNAATAGATTTATTTACTTNGAGTATGAATATTAGATTCAACGATAATGCCATAGTAGT  120
               || |||||||||||||||| |||||||||||||||||||||||||
||||||||||||
Sbjct  728736  CT-AATAGATTTATTTACTT-GAGTATGAATATTAGATTCAACGAT--
TGCCATAGTAGT  728791

Query  121
TGAAATTGAATTCCTTGCANTGTATTGATAAACGGTAGATACCTCTCTCTCATAAGAGAG  180
               ||||||||||||  |||||
||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  728792  TGAAATTGAATT--TTGCA-
TGTATTGATAAACGGTAGATACCTCTCTCTCATAAGAGAG  728848

Query  181     CACTTCATTTTCAATGGCCTTATAGACAAAAGCGTTAAGTGCTAATCCTTGC  232
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  728849  CACTTCATTTTCAATGGCCTTATAGACAAAAGCGTTAAGTGCTAATCCTTGC  72890
```

Hd7-2(21) *tdhA*

SEQUENCE:

ATTTCGCCCAGTACCCGCCGCTAACTCACTAAGATAAGGACGATATGATCGATTTAATAGGTTAT
CGACAGTGAATTGAATGCGTAAATGTTTGATTTGTTTCGGTTGCCAAGTCGCGAAGATGCCGTGT
AAAGCATAACCGCTTGATTTAGGTAATGCCCAATGGCCTGCATCTTTATCTTGATCTGTAGGAGA
ACGATCTTGACGTCGAACAAACTCACTTTTCCATCCTAACGTTAAATTTGTATCAGGAATATGGCT
ACCTAAAGTAACAACGGCTTTGCGAGGTGGAATTTCTGCAATCCAGCTTGTACTGGCTAGCCAAG
GGTTTCTTGGTTGGATGCTAATCNGTTTTCCTATTGTGTGTGAATAGCTAAGATTGGTGAACCAAT
AAGATTGAATCATAATTCGCCTCTNAACTCAAATCCTTTAATTTGATAGCCGGGTAAATTGCGGT
AATTGCCGATTTTTTTGCTACAAACTTCATNNN

BLASTX:

gb|AF052977.1|AF052977 Haemophilus ducreyi ton-dependent heme receptor A (TdhA) gene, complete cds
Length=2912

Score = 867 bits (469), Expect = 0.0
Identities = 482/487 (98%), Gaps = 5/487 (1%)
Strand=Plus/Minus

```
Query  1     ATTTCGCCCAGTACCCGCCGCTAACTCACTAAGATAAGGACGATATGATCGATTTAATAG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2767  ATTTCGCCCAGTACCCGCCGCTAACTCACTAAGATAAGGACGATATGATCGATTTAATAG  2708

Query  61    GTTATCGACAGTGAATTGAATGCGTAAATGTTTGATTTGTTTCGGTTGCCAAGTCGCGAA  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2707  GTTATCGACAGTGAATTGAATGCGTAAATGTTTGATTTGTTTCGGTTGCCAAGTCGCGAA  2648

Query  121   GATGCCGTGTAAAGCATAACCGCTTGATTTAGGTAATGCCCAATGGCCTGCATCTTTATC  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2647  GATGCCGTGTAAAGCATAACCGCTTGATTTAGGTAATGCCCAATGGCCTGCATCTTTATC  2588

Query  181   TTGATCTGTAGGAGAACGATCTTGACGTCGAACAAACTCACTTTTCCATCCTAACGTTAA  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2587  TTGATCTGTAGGAGAACGATCTTGACGTCGAACAAACTCACTTTTCCATCCTAACGTTAA  2528

Query  241   ATTTGTATCAGGAATATGGCTACCTAAAGTAACAACGGCTTTGCGAGGTGGAATTTCTGC  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2527  ATTTGTATCAGGAATATGGCTACCTAAAGTAACAACGGCTTTGCGAGGTGGAATTTCTGC  2468
```

```
Query  301   AATCCAGCTTGTACTGGCTAGCCAAGGGTTTCTTGGTTGGATGCTAATCNGTTTTCCTAT
360

Sbjct  2467  AATCCAGCTTGTACTGGCTAGCCAAGGGTTTCTTGGT-G-ATGCTAATC-GTTTTCCTAT
2411

Query  361   TGTGTGTGAATAGCTAAGATTGGTGAACCAATAAGATTGAATCATAATTCGCCTCTNAAC
420

Sbjct  2410  TGTGTGTGAATAGCTAAGATTGGTGAACCAATAAG-TTGAATCATAATTCGCCTCT-AAC
2353

Query  421   TCAAATCCTTTAATTTGATAGCCGGGTAAATTGCGGTAATTGCCGAttttttttGCTACAA
480

Sbjct  2352  TCAAATCCTTTÄATTTGATAGCCGGGTAAATTGCGGTAATTGCCGATTTTTTTGCTACAA
2293

Query  481   ACTTCAT  487

Sbjct  2292  ACTTCAT  2286
```

Hd4-18(17) *dnaK*

SEQUENCE:

CAGCTTGATTACGTGTTTGAACAAGTTCTTCAAATTGACGATCTGATTCAGCATTAGCTTCTGCAT
CACGAACCATTTGTTCTACTTCTTCATCGCTTAAGCCTGAAGATGCTTGAATTTTAATTTGTTGTTC
TTTACCAGTATTTTTATCTTTTGCTGATACGTTGATAATACCATTTGCATCAATATCAAATGTCACT
TCGATTTGTGGCATACCACGAGGTGCAGGATTAATACCTTCAAGGTTAAATTGACCTAAAGATTT
ATTGTCTGCAGCACGTATTACGTTCACCTTGTAACACGTGAATGGTTACTNGCTGATTGGNTTATC
TTCTGCCGTTGAGAACACTTGCGATTTTTTCGTTGGGATAGTGGTATTTNTTTTCAATTAGGCTTG
GTCATTACTCCGCCCATTGTTTCAATACCTAACGATAATGGTGTTACATCTAATAAAAGTACATCT
GTTACATCACCCGCTAACNNNN

BLASTX:

gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
chaperone protein DnaK

Score = 856 bits (463), Expect = 0.0
Identities = 476/481 (98%), Gaps = 5/481 (1%)
Strand=Plus/Plus

```
Query  1
CAGCTTGATTACGTGTTTGAACAAGTTCTTCAAATTGACGATCTGATTCAGCATTAGCTT  60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132576
CAGCTTGATTACGTGTTTGAACAAGTTCTTCAAATTGACGATCTGATTCAGCATTAGCTT  132635

Query  61
CTGCATCACGAACCATTTGTTCTACTTCTTCATCGCTTAAGCCTGAAGATGCTTGAATTT  120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132636
CTGCATCACGAACCATTTGTTCTACTTCTTCATCGCTTAAGCCTGAAGATGCTTGAATTT  132695

Query  121
TAATTTGTTGTTCTTTACCAGTATTTTTATCTTTTGCTGATACGTTGATAATACCATTTG  180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132696
TAATTTGTTGTTCTTTACCAGTATTTTTATCTTTTGCTGATACGTTGATAATACCATTTG  132755

Query  181
CATCAATATCAAATGTCACTTCGATTTGTGGCATACCACGAGGTGCAGGATTAATACCTT  240

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132756
CATCAATATCAAATGTCACTTCGATTTGTGGCATACCACGAGGTGCAGGATTAATACCTT  132815
```

```
Query  241
CAAGGTTAAATTGACCTAAAGATTTATTGTCTGCAGCACGTATTACGTTCACCTTGTAAC  300
                ||||||||||||||||||||||||||||||||||||||||||||||
||||||||||||||||||
Sbjct  132816  CAAGGTTAAATTGACCTAAAGATTTATTGTCTGCAGCACGT-
TTACGTTCACCTTGTAAC  132874

Query  301
ACGTGAATGGTTACTNGCTGATTGGNTTATCTTCTGCCGTTGAGAACACTTGCGATTTTT  360
                |||||||||||||| |||||||||
||||||||||||||||||||||||||||||||
Sbjct  132875  ACGTGAATGGTTACT-GCTGATTGG-
TTATCTTCTGCCGTTGAGAACACTTGCGATTTTT  132932

Query  361
TCGTTGGGATAGTGGTATTTNTTTTCAATTAGGCTTGGTCATTACTCCGCCCATTGTTTC  420
                |||||||||||||||||| ||||||||||||||||
||||||||||||||||||||
Sbjct  132933  TCGTTGGGATAGTGGTATTT-TTTTCAATTAGGCTTG-
TCATTACTCCGCCCATTGTTTC  132990

Query  421
AATACCTAACGATAATGGTGTTACATCTAATAAAAGTACATCTGTTACATCACCCGCTAA  480

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132991
AATACCTAACGATAATGGTGTTACATCTAATAAAAGTACATCTGTTACATCACCCGCTAA  133050

Query  481     C  481
                |
Sbjct  133051  C  133051
```

Hd12-19(17) *sodA*

SEQUENCE:

CGATTTTGATAATGTAAATAGTAAGCGTGTTCCCAAACATCTAAAGTAAAAATAGGGTAGCCTGA
AACGCCGGCTACATCTTTGCCCATAATAGTGCGAATCTTGGTTAGCAGTTGAAACAACCGCTAGT
TTGCCTTCTTCTAAGACCAGCNCAAGCCCATCCAGAACCAAAACGAGTCGCTGCGGCTTGTTCAA
ACTCGGCTTGGAATGCTTCTACGGAACCAAAATCGCGGATAATTGCATCTTNNNNN

BLASTX:

> gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
manganese superoxide dismutase

Score = 442 bits (239), Expect = 4e-121
Identities = 244/246 (99%), Gaps = 2/246 (0%)
Strand=Plus/Plus

```
Query  1
CGATTTTGATAATGTAAATAGTAAGCGTGTTCCCAAACATCTAAAGTAAAAATAGGGTAG  60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  256963
CGATTTTGATAATGTAAATAGTAAGCGTGTTCCCAAACATCTAAAGTAAAAATAGGGTAG  257022

Query  61
CCTGAAACGCCGGCTACATCTTTGCCCATAATAGTGCGAATCTTGGTTAGCAGTTGAAAC  120
                 ||||||||||||||||||||||||||||||||
||||||||||||||||||||||||
Sbjct  257023  CCTGAAACGCCGGCTACATCTTTGCCCATAATAG-
GCGAATCTTGGTTAGCAGTTGAAAC  257081

Query  121
AACCGCTAGTTTGCCTTCTTCTAAGACCAGCNCAAGCCCATCCAGAACCAAAACGAGTCG  180
                 ||||||||||||||||||||||||||||||||
||||||||||||||||||||||||
Sbjct  257082  AACCGCTAGTTTGCCTTCTTCTAAGACCAGC-
CAAGCCCATCCAGAACCAAAACGAGTCG  257140

Query  181
CTGCGGCTTGTTCAAACTCGGCTTGGAATGCTTCTACGGAACCAAAATCGCGGATAATTG  240

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  257141
CTGCGGCTTGTTCAAACTCGGCTTGGAATGCTTCTACGGAACCAAAATCGCGGATAATTG  257200

Query  241     CATCTT  246
                ||||||
Sbjct  257201  CATCTT  257206
```

Hd6-8(26) *cdtA*

SEQUENCE:

AACCTCAACCCCTATTATCAAAAACACCTTCAATGTCACTGAATTTGCTATCTTCATCCGGACCGA
ATAGACAGGTATTGCCGTCTGAACCATCAAACTTTATGACTTTGATGGGACAAAATGGGGCACTG
TTGACTGTCTGGGCGCTAGCAAAACGCA

BLASTX:

```
>  gb|AE017143.1|  Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

 Features in this part of subject sequence:
   cytolethal distending toxin protein A

 Score =  294 bits (159),  Expect = 6e-77
 Identities = 159/159 (100%), Gaps = 0/159 (0%)
 Strand=Plus/Plus

Query  1
AACCTCAACCCCTATTATCAAAAACACCTTCAATGTCACTGAATTTGCTATCTTCATCCG  60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  723369
AACCTCAACCCCTATTATCAAAAACACCTTCAATGTCACTGAATTTGCTATCTTCATCCG  723428

Query  61
GACCGAATAGACAGGTATTGCCGTCTGAACCATCAAACTTTATGACTTTGATGGGACAAA  120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  723429
GACCGAATAGACAGGTATTGCCGTCTGAACCATCAAACTTTATGACTTTGATGGGACAAA  723488

Query  121     ATGGGGCACTGTTGACTGTCTGGGCGCTAGCAAAACGCA  159
                ||||||||||||||||||||||||||||||||||||||
Sbjct  723489  ATGGGGCACTGTTGACTGTCTGGGCGCTAGCAAAACGCA  723527
```

Hd6-31(16) *norM*

SEQUENCE:

TGTGCCAGAAATGGGCGCCGTGGGCTGTGGCGTTGCGACCGCGATTGGTAATTGGTCGATGTTTT
TGATGATGTTACATTATTGCTATACGAATAAACCACAAAAAGACATTGGCTTATTTAACAAATGG
TTTGAAATGCCAAGCGGTAAGACTTTACTCAAAATTTGCAAATTAGGGCTACCTATTGGTTTTGCG
ACTTTTACTGA

BLASTX:

Distance tree of results NEW

Sequences producing significant alignments:	Score (Bits)	E Value
gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete ge...	372	4e-100

Alignments

> gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
Multidrug resistance protein NorM

Score = 372 bits (201), Expect = 4e-100
Identities = 205/207 (99%), Gaps = 0/207 (0%)
Strand=Plus/Plus

```
Query  1
TGTGCCAGAAATGGGCGCCGTGGGCTGTGGCGTTGCGACCGCGATTGGTAATTGGTCGAT  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
||||||||||||
Sbjct  854280
TGTGCCAGAAATGGGCGCCGTGGGCTGTGGCGTTGCGACCGCGATTGTTAATTGGTCGAT  854339

Query  61
GTTTTTGATGATGTTACATTATTGCTATACGAATAAACCACAAAAAGACATTGGCTTATT  120
               ||||||||||||||||
||||||||||||||||||||||||||||||||||||||||||
Sbjct  854340
GTTTTTGATGATGTTTCATTATTGCTATACGAATAAACCACAAAAAGACATTGGCTTATT  854399
```

```
Query  121
TAACAAATGGTTTGAAATGCCAAGCGGTAAGACTTTACTCAAAATTTGCAAATTAGGGCT  180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  854400
TAACAAATGGTTTGAAATGCCAAGCGGTAAGACTTTACTCAAAATTTGCAAATTAGGGCT  854459

Query  181      ACCTATTGGTTTTGCGACTTTTACTGA  207
                 ||||||||||||||||||||||||||
Sbjct  854460  ACCTATTGGTTTTGCGACTTTTACTGA  854486
```

Hd9-13(20) *uspA*

SEQUENCE:

ATATTCTAGTTGCCAGTTGATCTTTCTGAAGAAAGTTTGATTTTACTTCGTAAAGGTGCGGGATTG
GCTGAGAAAATGTGGTGCTAAACCTTTCACTTATTGCATGTAGATGTCAATTTTTCTGATCTTTAT
ACCGGATTAATTGATA

BLASTX:

Distance tree of results NEW

Score
E
Sequences producing significant alignments: (Bits)
Value

gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete ge... 248
5e-63

Alignments

> gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
universal stress protein A

Score = 248 bits (134), Expect = 5e-63
Identities = 144/148 (97%), Gaps = 4/148 (2%)
Strand=Plus/Plus

```
Query  1
ATATTCTAGTTGCCAGTTGATCTTTCTGAAGAAAGTTTGATTTTACTTCGTAAAGGTGCG  60
                  ||||||||||||
||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1169142  ATATTCTAGTTG-
CAGTTGATCTTTCTGAAGAAAGTTTGATTTTACTTCGTAAAGGTGCG  1169200

Query  61
GGATTGGCTGAGAAAATGTGGTGCTAAACCTTTCACTTATTGCATGTAGATGTCAATTTT  120
                  |||||||||||||||| |||||||||||| ||||||||||
||||||||||||||||||
Sbjct  1169201  GGATTGGCTGAGAAA-TGTGGTGCTAAAC-TTTCACTTATT-
CATGTAGATGTCAATTTT  1169257

Query  121      TCTGATCTTTATACCGGATTAATTGATA  148
                 ||||||||||||||||||||||||||||
Sbjct  1169258  TCTGATCTTTATACCGGATTAATTGATA  1169285
```

Hd1-1(24) *mutS*

SEQUENCE:

ATCTAATTGCCTAATCCGGTTGATGAATCCACCGCTATTAATAAACGGCTGCCCATTGGCGTAAC
GCATTTATCTAATACCGAAGCAAGCGTTGCTTCTGTACCGCGCCGCTAAATTTTGAGTTAGCTCTA
AATTTCTACGGGTGGCGGCATCTAATAACACCATTTCACTATCTTGCGAGAGATGAATGCTATTA
AAA

BLASTX:

```
>  gb|AE017143.1|  Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

 Features in this part of subject sequence:
   DNA mismatch repair protein MutS

 Score =  344 bits (186),  Expect = 8e-92
 Identities = 194/197 (98%), Gaps = 3/197 (1%)
 Strand=Plus/Plus

Query  1
ATCTAATTGCCTAATCCGGTTGATGAATCCACCGCTATTAATAAACGGCTGCCCATTGGC  60
                |||||||||||||| ||||||||||||||||||||
||||||||||||||||||||||
Sbjct  1685616  ATCTAATTGCCTAAT-CGGTTGATGAATCCACCGCT-
TTAATAAACGGCTGCCCATTGGC  1685673

Query  61
GTAACGCATTTATCTAATACCGAAGCAAGCGTTGCTTCTGTACCGCGCCGCTAAATTTTG  120
                |||||||||||||||||||||||||||||||||||||||||||||||||||
||||||||||||
Sbjct  1685674  GTAACGCATTTATCTAATACCGAAGCAAGCGTTGCTTCTGTACCGC-
CCGCTAAATTTTG  1685732

Query  121
AGTTAGCTCTAAATTTCTACGGGTGGCGGCATCTAATAACACCATTTCACTATCTTGCGA  180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1685733
AGTTAGCTCTAAATTTCTACGGGTGGCGGCATCTAATAACACCATTTCACTATCTTGCGA  1685792

Query  181      GAGATGAATGCTATTAA  197
                |||||||||||||||||
Sbjct  1685793  GAGATGAATGCTATTAA  1685809
```

Hd8-6(16) *hgbA*

SEQUENCE:

AGCAGGACAATCAATTTCACCATTAACGTCACAATATTCATCTAAACGAGCTTTCAATTTAATTTT
CTGGTGATTATGGGTAATTTTGAGTGTATCCCAAAAGTCATTAGTATCAAAATTTTCATAGACAA
AAGAAATATTTTTACGATTACTCTGATCATTAGTATGGCGTACATCTTTTAAAAATGGCCCACCAT
TATAAGAGGTAAATGCATTAGACCAATCA

BLASTX:

```
>  gb|AE017143.1|  Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

 Features in this part of subject sequence:
   hemoglobin-binding protein HgbA

 Score =  418 bits (226),  Expect = 5e-114
 Identities = 226/226 (100%), Gaps = 0/226 (0%)
 Strand=Plus/Plus

Query  1
AGCAGGACAATCAATTTCACCATTAACGTCACAATATTCATCTAAACGAGCTTTCAATTT  60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1688560
AGCAGGACAATCAATTTCACCATTAACGTCACAATATTCATCTAAACGAGCTTTCAATTT  1688619

Query  61
AATTTTCTGGTGATTATGGGTAATTTTGAGTGTATCCCAAAAGTCATTAGTATCAAAATT  120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1688620
AATTTTCTGGTGATTATGGGTAATTTTGAGTGTATCCCAAAAGTCATTAGTATCAAAATT  1688679

Query  121
TTCATAGACAAAAGAAATATTTTTACGATTACTCTGATCATTAGTATGGCGTACATCTTT  180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1688680
TTCATAGACAAAAGAAATATTTTTACGATTACTCTGATCATTAGTATGGCGTACATCTTT  1688739

Query  181      TAAAAATGGCCCACCATTATAAGAGGTAAATGCATTAGACCAATCA  226
                ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1688740  TAAAAATGGCCCACCATTATAAGAGGTAAATGCATTAGACCAATCA  1688785
```

Hd6-34(16) *hlyX*

SEQUENCE:

TTGTACGGTAATCATACCACTTTTCTGGAAGCGACCCACAAGCAAACGGCTAATGGTTTCTATCG
TTAAGCCTAAATAATTACCGATATCGCCACGTGTCATTGTCACAACGGAATTCACGTGCAGAGAA
A

BLASTX:

```
Sequences producing significant alignments:                          (Bits)
Value

gb|AE017143.1|  Haemophilus ducreyi strain 35000HP complete ge...   217
1e-53
```

Alignments

```
>  gb|AE017143.1|  Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

 Features in this part of subject sequence:
   fnr-like transcriptional regulator protein

 Score =  217 bits (117),  Expect = 1e-53
 Identities = 127/131 (96%), Gaps = 4/131 (3%)
 Strand=Plus/Plus

Query  1
TTGTACGGTAATCATACCACTTTTCTGGAAGCGACCCACAAGCAAACGGCTAATGGTTTC  60
               ||||||||||||||||||||||||||||||  |
||||||||||||||||||||
Sbjct  1168202  TTGTACGGTAATCATACCACTTTTCTGGAAGCGA--C-
CAAGCAAACGGCTAATGGTTTC  1168258

Query  61
TATCGTTAAGCCTAAATAATTACCGATATCGCCACGTGTCATTGTCACAACGGAATTCAC  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
||||||||||||
Sbjct  1168259  TATCGTTAAGCCTAAATAATTACCGATATCGCCACGTGTCATTGTCA-
AACGGAATTCAC  1168317

Query  121      GTGCAGAGAAA  131
                ||||||||||||
Sbjct  1168318  GTGCAGAGAAA  1168328
```

Hd7-5(13) *greA*

SEQUENCE:

TTCTACTTTACCACCCGGGGTAATAATACTGACTGAATCATCGACACTTTTACCCACTAAACCGCG
CGCAATCGGCGAATTAACTGAGATTAATCCTATCTTTAATATTGGCCTCATCATCACCAACAATAC
GATAAGTCACTTCTTCATCTGTATCAAATATTCACTAGCCCAACAGTTGCACCAAAAATCACTTTA
CCATTATTGGTCATTTTTGCTACATCAATCACTTGTGCATTGCCTAATTTTCCTTCAATCCA

BLASTX:

gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
transcription elongation factor

Score = 462 bits (250), Expect = 3e-127
Identities = 255/257 (99%), Gaps = 2/257 (0%)
Strand=Plus/Plus

```
Query  1
TTCTACTTTACCACCCGGGGTAATAATACTGACTGAATCATCGACACTTTTACCCACTAA  60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1415070
TTCTACTTTACCACCCGGGGTAATAATACTGACTGAATCATCGACACTTTTACCCACTAA  1415129

Query  61
ACCGCGCGCAATCGGCGAATTAACTGAGATTAATCCTATCTTTAATATTGGCCTCATCAT  120
                  ||||||||||||||||||||||||||||||
||||||||||||||||||||
Sbjct  1415130  ACCGCGCGCAATCGGCGAATTAACTGAGATTAATCCT-
TCTTTAATATTGGCCTCATCAT  1415188

Query  121
CACCAACAATACGATAAGTCACTTCTTCATCTGTATCAAATATTCACTAGCCCAACAGTT  180
                  ||||||||||||||||||||||||||||||||||
||||||||||||||||||||
Sbjct  1415189  CACCAACAATACGATAAGTCACTTCTTCATCTGTATC-
AATATTCACTAGCCCAACAGTT  1415247

Query  181
GCACCAAAAATCACTTTACCATTATTGGTCATTTTTGCTACATCAATCACTTGTGCATTG  240

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1415248
GCACCAAAAATCACTTTACCATTATTGGTCATTTTTGCTACATCAATCACTTGTGCATTG  1415307

Query  241      CCTAATTTTCCTTCAAT  257
                 |||||||||||||||||
Sbjct  1415308  CCTAATTTTCCTTCAAT  1415324
```

gb|CP000057.2| Haemophilus influenzae 86-028NP, complete genome
Length=1914490

Features in this part of subject sequence:
transcription elongation factor GreA

Score = 171 bits (92), Expect = 2e-39
Identities = 201/253 (79%), Gaps = 10/253 (3%)
Strand=Plus/Minus

```
Query  10       ACCACCCGGGGTAATAATACTGACTGAATCATCGACACT-
TTTACCCACTAAACCGCGCG  68
                ||||||  ||  ||  |||||  |  ||  |  ||||||  |  |  |  ||||||  |  |||  ||
||||
Sbjct  1538868  ACCACCAGGCGTTGTAATATTAACGGTATCATCTA-
ATTCTTTACCAATTAAGCCTCGCG  1538810

Query  69
CAATCGGCGAATTAACTGAGATTAATCCTATCTTTAATATTGGCCTCATCATCACCAACA  128
                ||||  |||||  ||  ||  ||  |||||  ||  |  ||||||||  ||||||||||  ||
|||
Sbjct  1538809  CAATGGGCGAGTTTACCGAAATTAAACC-
AGATTTAATATCAGCCTCATCATCGCCTACA  1538751

Query  129
ATACGATAAGTCACTTCTTCATCTGTATCAAATATTCACTAGCCCAACAGTTGCACCAAA  188
                ||  ||||||||  ||||||||  |||||||    ||||  ||  ||  ||||
||||||||||
Sbjct  1538750  ATGCGATAAGTGACTTCTTCGTCTGTAT-
TTGTATTTACAAGTACAACTGTTGCACCAAA  1538692

Query  189      AATCACTTTACCATTATT-GGTCATTTTTGCT-
ACATCAATCACTTGTGCATTGCCTAAT  246
                ||||||||||||||||||  ||  ||||||||  |  ||||||||  |  |||  |||||  ||
||
Sbjct  1538691  AATCACTTTACCATTATTTGG-CATTTTTG-TGACATCAATAATTTGAGCATTACC-
AAG  1538635

Query  247      TTT-CCTTCAATC  258
                |||  |||||||||
Sbjct  1538634  TTTACCTTCAATC  1538622
```

Hd3-2(26) *ptrA*

SEQUENCE:

TACTTGAAGCATTGGCACGTAAAAATAAAGCCAATAGTTTAGAACAAACGAATCGTATCCTATCA
AATTTTGCTGATTATCCTTATTTTGAAGAAGATAAACAGCGTAAAGTGATTAATGAGATTACTTTA
GCAGATATTAAAGCGATACGTGAAAAAGTATTGAGTAAACCAACCGGTGTAAATGTGTTATCTGT
GGGGAATTTAGATGATTCACAGGTTAAACAACTTGTGCAACACGTGAATGAAGTGGTACAGAAT
CGCAATCTTGAGTTAGGTAAAGCACGTTATCTTCGATTTAAATGAGAGCGAGCGCAAGTTTAATT
ATGTACAAACCGTCGCCTCATGAGGATAATGCCGTTAAATGTAACTTATTTAGCTAAAGGCTATG
ATGAAGTTAGATAGTTCTATCCGATCGCCGGTTATTAAGTGATATTATTAGCCGTTGGTAA

BLASTX:

Score
E
Sequences producing significant alignments: (Bits)
Value

gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete ge... 798
0.0

Alignments

>gb|AE017143.1| Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

Features in this part of subject sequence:
protease III

Score = 798 bits (432), Expect = 0.0
Identities = 445/450 (98%), Gaps = 5/450 (1%)
Strand=Plus/Plus

```
Query  1
TACTTGAAGCATTGGCACGTAAAAATAAAGCCAATAGTTTAGAACAAACGAATCGTATCC  60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  129864
TACTTGAAGCATTGGCACGTAAAAATAAAGCCAATAGTTTAGAACAAACGAATCGTATCC  129923

Query  61
TATCAAATTTTGCTGATTATCCTTATTTTGAAGAAGATAAACAGCGTAAAGTGATTAATG  120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  129924
TATCAAATTTTGCTGATTATCCTTATTTTGAAGAAGATAAACAGCGTAAAGTGATTAATG  129983

Query  121
AGATTACTTTAGCAGATATTAAAGCGATACGTGAAAAAGTATTGAGTAAACCAACCGGTG  180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  129984
AGATTACTTTAGCAGATATTAAAGCGATACGTGAAAAAGTATTGAGTAAACCAACCGGTG  130043

Query  181
TAAATGTGTTATCTGTGGGGAATTTAGATGATTCACAGGTTAAACAACTTGTGCAACACG  240

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  130044
TAAATGTGTTATCTGTGGGGAATTTAGATGATTCACAGGTTAAACAACTTGTGCAACACG  130103

Query  241
TGAATGAAGTGGTACAGAATCGCAATCTTGAGTTAGGTAAAGCACGTTATCTTCGATTTA  300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
||||||
Sbjct  130104  TGAATGAAGTGGTACAGAATCGCAATCTTGAGTTAGGTAAAGCACGTTATCTT-
GATTTA  130162

Query  301
AATGAGAGCGAGCGCAAGTTTAATTATGTACAAACCGTCGCCTCATGAGGATAATGCCGT  360
                ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
||
Sbjct  130163  AATGAGAGCGAGCGCAAGTTTAATTATGTACAAACCGTC-CCTCATGAGGATAATGC-
GT  130220

Query  361
TAAATGTAACTTATTTAGCTAAAGGCTATGATGAAGTTAGATAGTTCTATCCGATCGCCG  420
                ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
|
Sbjct  130221  TAAATGTAACTTATTTAGCTAAAGGCTATGATGA-GTTAGATAGTTCTATCCGATCGC-
G  130278

Query  421     GTTATTAAGTGATATTATTAGCCGTTGGTA  450
                ||||||||||||||||||||||||||||
Sbjct  130279  GTTATTAAGTGATATTATTAGCCGTTGGTA  130308
```

Hd5-23(16) *radA*

SEQUENCE:

GCCAAAATATCTAAAGTTTGACTTAATGTTTTAATCGCAAAAATTCCCATTCCTTGAAAATCGTTT
TTTTCGGCATATTGCCGTATGGAATAATGGGCGCGTCGAAAACCGTGCTTAAGTGGCTTCACTCA
TTCTTT

BLASTX:

Distance tree of results NEW

```
                                                                    Score
E
Sequences producing significant alignments:                         (Bits)
Value

gb|AE017143.1|  Haemophilus ducreyi strain 35000HP complete ge...   217
1e-53
```

Alignments

```
>   gb|AE017143.1|  Haemophilus ducreyi strain 35000HP complete genome
Length=1698955

 Features in this part of subject sequence:
   DNA repair protein RadA

 Score =  217 bits (117),  Expect = 1e-53
 Identities = 131/137 (95%),  Gaps = 4/137 (2%)
 Strand=Plus/Plus

Query  1
GCCAAAATATCTAAAGTTTGACTTAATGTTTTAATCGCAAAAATTCCCATTCCTTGAAAA  60
              |||||||||||||||||||||||||||||||||| |||||||||| ||||||||||||
Sbjct  890952  GCCAAAATATCTAAAGTTTGACTTAATGTTTTAACCGCAAAAATTTCCATTCCTTGAA-
-  891009

Query  61
TCGttttttttCGGCATATTGCCGTATGGAATAATGGGCGCGTCGAAAACCGTGCTTAAGT  120
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||
||
Sbjct  891010  TCGTTTTTTTTCGGCATATTGCCGTATGGAATAATGG-CGCGTCGAAAACCGTGCTTA-
GT  891067

Query  121     GGCTTCACTCATTCTTT  137
               |||||||||||||||||
Sbjct  891068  GGCTTCACTCATTCTTT  891084
```