Investigating the Role of Pallilysin in the Dissemination of the Syphilis Spirochete *Treponema pallidum*

by

Yavor Denchev
Bachelor of Science, University of Victoria, 2011

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Biochemistry and Microbiology

Supervisory Committee

Investigating the Role of Pallilysin in the Dissemination of the Syphilis Spirochete *Treponema pallidum*

by

Yavor Denchev
Bachelor of Science, University of Victoria, 2011

Supervisory Committee

Dr. Caroline Cameron, Department of Biochemistry and Microbiology
Supervisor

Dr. Terry Pearson, Department of Biochemistry and Microbiology
Departmental Member

Dr. John Taylor, Department of Biology
Outside Member

Abstract

Supervisory Committee
Dr. Caroline Cameron, Department of Biochemistry and Microbiology
Supervisor
Dr. Terry Pearson, Department of Biochemistry and Microbiology
Departmental Member
Dr. John Taylor, Department of Biology
Outside Member

Syphilis is a global public health concern with 36.4 million cases worldwide and 11 million new infections per year. It is a chronic multistage disease caused by the spirochete bacterium *Treponema pallidum* and is transmitted by sexual contact, direct contact with lesions or vertically from an infected mother to her fetus. *T. pallidum* is a highly invasive pathogen that rapidly penetrates tight junctions of endothelial cells and disseminates rapidly via the bloodstream to establish widespread infection. Previous investigations conducted in our laboratory identified the surface-exposed adhesin, pallilysin, as a metalloprotease that degrades the host components laminin (major component of the basement membrane lining blood vessels) and fibrinogen (primary component of the coagulation cascade), as well as fibrin clots (function to entrap bacteria and prevent disseminated infection). Furthermore, pallilysin expressed on the surface of the non-invasive spirochete *Treponema phagedenis* conferred upon this bacterium the ability to degrade fibrin clots. It was hypothesized that pallilysin is integral to the process of *T. pallidum* dissemination, and interference with its functioning will prevent spread throughout the host and establishment of chronic infection. To test this hypothesis, a two-pronged approach was undertaken during my thesis research.

Bioinformatics analyses were used to trace the evolutionary history of pallilysin in an attempt to gain further insight into its role in the pathogenesis of *T. pallidum*. The

sequence conservation of pallilysin was analyzed in the context of its homologues. The bioinformatics analyses revealed homologues in three spirochete genera, namely *Treponema*, *Spirochaeta*, and *Borrelia*, presented in decreasing order of the degree of sequence conservation. The HEXXH motif, part of the active site of the pallilysin metalloprotease, was fully conserved only in *T. pallidum* and *T. paraluiscuniculi*, both of which are systemic pathogens. However, the flanking sequences showed a high degree of conservation, especially in the *Treponema* and *Spirochaeta* genera. The minimum laminin-binding region of pallilysin identified previously was partially conserved among the treponema and spirochaeta homologues with the highest degree of conservation observed with the homologues from *T. paraluiscuniculi* and *T. phagedenis*, as well as among the homologues from the human oral pathogens.

In vitro dissemination studies were performed to investigate the dissemination capacity of *T. phagedenis* heterologously expressing pallilysin. Human Umbilical Vein Endothelial Cells were seeded and grown to confluence on permeable inserts coated with growth factor-reduced Matrigel to create an artificial endothelial barrier. Wild type *T. phagedenis*, and *T. phagedenis* transformed either with the pallilysin open reading frame or its empty shuttle vector, were incubated with the barriers under anaerobic conditions. Dissemination across the barrier was assessed as percent traversal by both dark-field microscopic counts of treponemes and real-time quantitative PCR of genomic DNA extracted from the treponemes. The results were inconclusive. However, a traversal trend suggested heterologous expression of pallilysin may facilitate traversal of *T. phagedenis* across the artificial endothelial barrier.

This study presented the first step towards elucidating the role of pallilysin in endothelial monolayer traversal and provided supporting evidence for the role of pallilysin in the widespread dissemination of *T. pallidum in vivo*.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Sexually transmitted infections	(STIs)
World Health Organization	(WHO)
Public Health Agency of Canada	(PHAC)
Men who have sex with men	(MSM)
Human immunodeficiency virus	(HIV)
Mega base pairs	(Mb)
Degrees Celsius	(°C)
Lipopolysaccharide	(LPS)
Outer membrane	(OM)
Lipoprotein	(LP)
Peptidoglycan	(PG)
Cell (inner) membrane	(CM)
Outer membrane proteins	(OMP)
Cerebrospinal fluid	(CSF)
Extracellular matrix	(ECM)
Matrix metalloproteinase-1	(MMP-1)
Intercellular adhesion molecule 1	(ICAM-1)
Transendothelial electrical resistance	(TEER)
Shuttle vector pKMR4PE	(pKMR)
Erythromycin resistance	(emr)
Human Umbillical Vein Endothelial Cells	(HUVEC)
Real-time quantitative PCR	(QPCR)
Tryptone-yeast extract-gelatin-volatile fatty acids-serum	(TYGVS)
Heat-inactivated normal rabbit serum	(HI-NRS)
Heat-inactivated fetal bovine serum	(HI-FBS)
Biosafety Cabinet Class II	(BSCII)
Trypsin Neutralizing Solution	(TNS)
Polyethylene terephthalate	(PET)

Polycarbonate	(PC)
Engelbreth-Holm-Swarm	(EHS)
Growth factor-reduced	(GFR)
Dissemination Assay Mixed Media	(DAM media)
flagellar hook polypeptide gene	*(flgE)*
SDS polyacrylamide gel electrophoresis	(SDS-PAGE)
Microbial Genome Database	(MBGD)
National Center for Biotechnology Information	(NCBI)
Basic Local Alignment Search Tool	(BLASTp)
Position-Specific Iterated BLAST	(PSI-BLAST)
Open Reading Frame	(ORF)
Hypothetical proteins	(HP)
Profile hidden Markov models	(profile HMMs)
Hypoxia Inducible Factor	(HIF)
Vascular endothelial growth factor	(VEGF)
Melting temperature	(Tm)
Electric cell substrate impedance sensing	(ECIS)
3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide	(MTT)
Lactate Dehydrogenase	(LDH)

Acknowledgments

First of all, I would like to thank my supervisor, Dr. Caroline Cameron, for the invaluable instruction and motivation in the past three years of my graduate studies. In addition, I would like to acknowledge her support of my ideas in letting me take ownership of my research project. Finally, I would like to thank her for the understanding of my limits of comfort in working on research that involves animal models and honoring my request to not work on such. I hold the highest admiration for her dedication to science, which has been an inspiration for me to keep pushing my own limits.

Secondly, my committee members were always available for discussion or instruction. I would like to thank Dr. Terry Pearson for allowing me to occupy a significant amount of space in his laboratory by using his Biosafety cabinet. Apart from providing the equipment crucial for my project, he and his lab made it a very welcoming place for me. Dr. John Taylor was kind enough to teach me about bioinformatics and fuel my excitement about my project even further. Together with my supervisor, my committee members formed the dream team to instruct and motivate me to complete my research thesis.

My labmates have been the most incredible, supportive group of people I have ever worked with and this would not have been possible without them. From personal to professional, they have been there to back me when in need. Charmaine Wetherell and Rebecca Hoff were instrumental in developing and carrying out of my research project and I cannot thank them enough. I would also like to acknowledge Drs. Simon Houston and Tim Witchel for the many fruitful discussions and help with my project. Finally, to all the rest of the lab members – you are amazing and you are an inspiration!

Last but not least, great many thanks to Albert Labossiere, Stephen Horak, and Scott Scholz for building an air-conditioned chamber around the anaerobic chamber, without which I would not have been able to finish my experiments.

I swore to acknowledge my lab coat if my experiment ever worked, so huge thanks for that. And last but not least, I would like to debunk a common misconception about using voodoo dolls in science to positively affect the experimental outcome. Although in certain fields that might prove useful, *treponeme* research seems to be immune to the influence of *treponeme* effigies. Or perhaps, paper was the wrong material...

Dedication

I would like to dedicate this work to my fiancée, family and friends for their continuing support in all my endeavours. You kept me on track and this is for you!

And perhaps not so much a dedication, but a note to myself– sometimes it is not impossible to make the impossible possible – keep persevering!

Chapter 1: Introduction

1.1 Syphilis

Syphilis is a chronic multistage disease, caused by the bacterium, *Treponema pallidum* subsp. *pallidum* (*T. pallidum*) (Lafond and Lukehart, 2006). Humans are the only known natural hosts for *T. pallidum.* The bacteria are transmitted via sexual contact, or direct contact with primary and secondary lesions. Infection can also occur via vertical transmission of the pathogen from an infected mother to her fetus *in utero.* Although painless in the early stages, untreated syphilis can progress with neurological, cardiovascular and skeletal complications later in life, as well as result in fetal loss in pregnant women with acute infection (World Health Organization, 2012).

1.1.1 Epidemiology

Sexually transmitted infections (STIs) are a global public health concern responsible for acute illness, infertility, long-term disability, and death with dire physical and psychological complications for a myriad of men, women and infants (World Health Organization, 2012). Out of 30 identified bacterial, viral and parasitic pathogens transmitted sexually, it was estimated that in 2008 there were 100.4 million adults infected with *C. trachomatis*, 36.4 million with *N. gonorrhoeae*, 36.4 million with syphilis, and 187.0 million with *T. vaginalis* (World Health Organization, 2012). A preventable disease, infectious syphilis remains a critical public health burden with 10.6 million new infections per year. Although the World Health Organization (WHO) estimated that the vast majority of syphilis cases were found in developing nations, the

disease has resurged recently in developed countries, such as Canada (Health, 2007), the US (Lancet, 2011), the UK (Simms *et al*., 2005), China (Hesketh *et al*., 2010), and within Europe (Herbert and Middleton, 2012).

Congenital syphilis remains a major public health concern associated with the highest rate of fetal loss or stillbirth in low income settings (Goldenberg, 2003; Watson-Jones *et al*., 2002). While cases of congenital syphilis infections predominate in sub-Saharan Africa (Lozano *et al*., 2012), where lack of prenatal testing and antibiotic treatment of infected pregnant women are common (Lafond and Lukehart, 2006), developed countries, such as China (Hesketh *et al*., 2010) and Canada (Singh *et al*., 2007) are faced with a rising health burden of congenital syphilis. In 2008, the WHO estimated that globally each year 1.86 million pregnant women are infected with syphilis, with 4% to 15% of pregnant women in Africa being infected (World Health Organization, 2007), and that a substantial proportion of these individuals do not get treated adequately or at all (Gomez *et al*., 2013). Up to one third of the women attending antenatal care clinics are not tested for syphilis (World Health Organization, 2011). The consequences of untreated early syphilis in pregnant women include stillbirth (25%) and neonatal death (14%) with an overall perinatal mortality of close to 40% (World Health Organization, 2007).

Syphilis has been nationally notifiable in Canada since 1924 (Public Health Agency of Canada, 2008). From 1993 to 2000, reported rates of infectious syphilis were relatively stable (0.6 per 100, 000) and similar between genders (Figure 1). Reported rates began to increase sharply in 2001, more pronounced in men than in women. In 2008, 1,394 cases of infectious syphilis were reported to the Public Health Agency of Canada (PHAC), resulting in a rate of 4.2 per 100,000. Between 1999 and 2008, the rate

in men increased from 0.7 to 7.3 per 100,000 and in women increased from 0.5 to 1.1 per 100,000.

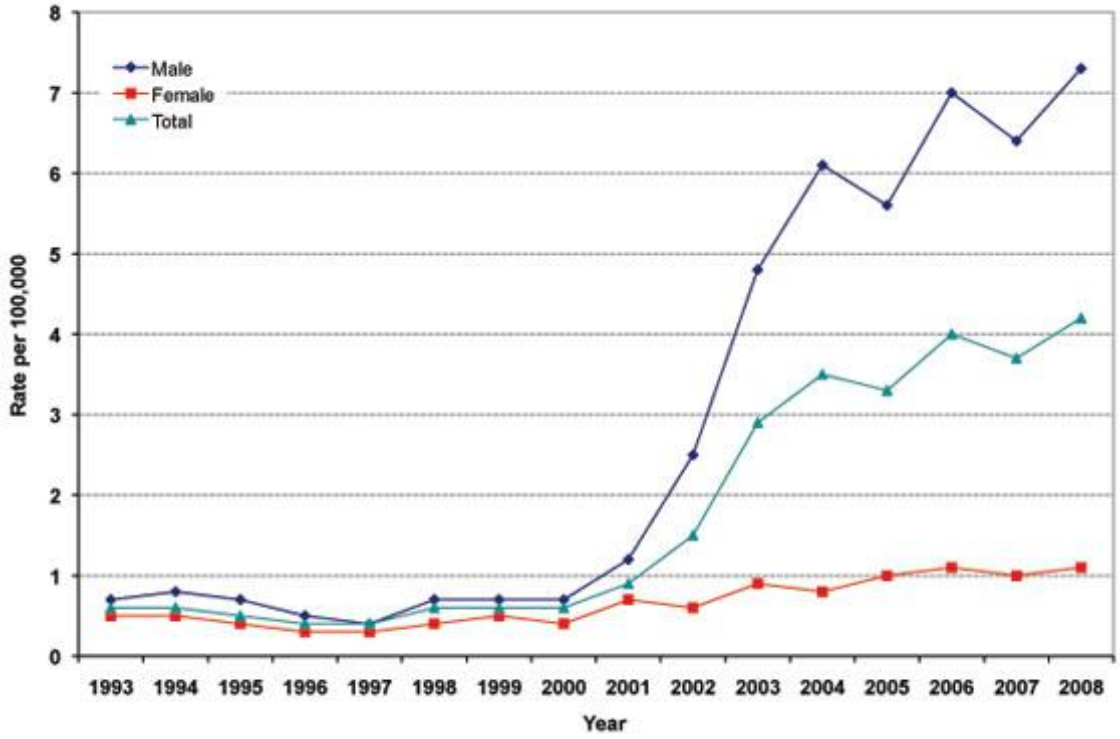


Figure 1. Reported rates of infectious syphilis by sex and overall, 1993 to 2008, Canada (Public Health Agency of Canada, 2008)

Within Canada, the rates of infectious syphilis in British Columbia (BC) have increased over the past fifteen years (BC Centre for Disease Control, 2013). Infectious syphilis rates started to increase in 1997 from a decade long stable rate of below 2 per 100,000. The increase was initially among vulnerable street populations, such as sex trade workers and people with unstable housing. Since 2004 the cases were primarily among bisexual and men who have sex with men (MSM) populations with more than half of them also being HIV-positive. By 2006, BC reached a rate nearly double that of Canada (7.8 cases per 100,000 population as compared with 4.0 per 100,000) (Public

Health Agency of Canada, 2008). The rate continued to increase steadily, apart from a sporadic drop in infections syphilis cases in 2009-2010, and reached the highest rate in over 30 years in 2012 with 8 per 100,000 (BC Centre for Disease Control, 2013). The male population comprised over 90% (346 cases) of infectious syphilis cases in 2012, with the highest rates observed in males between 25-59 years of age (Communicable Disease Prevention BC CDC, 2013). Across BC the highest rates of infectious syphilis in 2012 were in the following health service delivery areas: Vancouver (37.5 per 100,000 populaton; 257 cases), Richmond (6.0 per 100,000 populaton; 12 cases) and Fraser North (5.0 per 100,000 populaton; 31 cases).

Further concern for public health is the fact that infectious syphilis increases the risk for acquisition and transmission of the human immunodeficiency virus (HIV) 2–5-fold (Buchacz *et al*., 2004; Nusbaum *et al*., 2004). Currently no preventative vaccine for syphilis exists, which underlines the necessity for a greater understanding of the molecular mechanisms of *T. pallidum* pathogenesis (Cameron and Lukehart, 2013b).

1.1.2. Syphilis Pathology and Course of Infection

Syphilis is a multistage disease that manifests with localized, disseminated, and chronic phases depending on the stage of infection (Ho and Lukehart, 2011). Clinically, the disease presents with both symptomatic and prolonged asymptomatic stages because *T. pallidum* is able to evade the host immune response and remain latent for long periods of time. If left untreated, syphilis results in a systemic infection with serious medical and psychological consequences. Damage can be incurred to any organ system, including bones, heart, aorta, eyes, and brain (Lafond and Lukehart, 2006).

Syphilis shows different clinical manifestations throughout the natural course of untreated infection. A schematic diagram of the consecutive stages of untreated syphilis is presented in Figure 2.

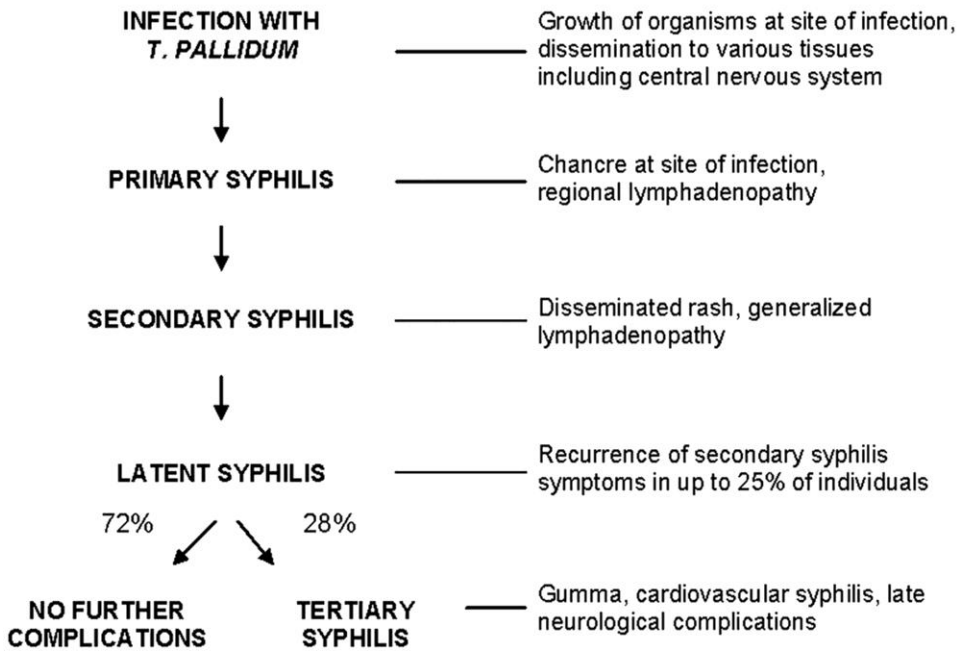


Figure 2. Schematic representation of the natural course of infection in untreated syphilis (Lafond and Lukehart, 2006)

The primary stage of infection begins when *T. pallidum* penetrates microscopic dermal abrasions or intact mucous membranes causing a local inflammatory response. The result is a painless chancre that develops approximately 3–6 weeks after the initial infection. The primary chancre usually occurs at the site of infection and becomes indurated followed by ulceration. Within 3–8 weeks, the chancre heals with or without treatment, indicating local clearance of *T. pallidum*. However, the stage for secondary

infection has been set, because by that time *T. pallidum* has already spread systemically to a multitude of tissues and organs. The painless nature and spontaneous healing of the primary chancre may leave infected people unaware in the face of a developing syphilis infection. The next stage is marked by a disseminated rash – a common manifestation of secondary syphilis that occurs usually within 3 months of infection. The rash is commonly found on the soles of the feet and the palms of the hands. While neurological complications were classically associated with the tertiary stage of the disease, *T. pallidum* is able to disseminate to and penetrate the central nervous system during earlier stages as well, accounting for 40% of early syphilis cases that show neurological complications. Both primary and secondary lesions are infectious by direct contact and transmission of the pathogen can also occur vertically mother-to-fetus. Even though the host immune system effectively clears *T. pallidum* from the local primary and secondary sites of infection, the pathogen remains in multiple tissues without causing any clinical manifestations. At that point syphilis has entered the asymptomatic latent stage; however, infection from a mother to her fetus can still occur and 25% of patients may show symptoms of secondary syphilis during the latent stage. If the latent stage lasts for longer than a year it is deemed late chronic infection. In the majority of individuals with chronic latent syphilis no further clinical symptoms are observed, although these individuals remain infected for their lifetime. However, a retrospective study in the pre-antibiotic era showed that about a third of infected patients develop tertiary syphilis years or decades after the initial infection (Gjestland, 1955). Tertiary syphilis affects a multitude of organs and causes clinical symptoms including gumma, cardiovascular complications, and neurosyphilis. Currently, a vaccine for syphilis is not available, and the suggested

treatment for all stages of the disease is penicillin, in the form of penicillin G (Centers for Disease Control and Prevention, 2006).

Syphilis increases the risk for transmission and acquisition of HIV due to the disruption of epithelial and mucosal barriers at the primary chancre (Ghanem *et al*., 2009). Additionally, the increased number of immune cells attracted to the syphilis lesions facilitates infection with HIV. Moreover, both *T. pallidum* and its lipoproteins increase the expression of a chemokine receptor on macrophages and dendritic cells, which also serves as a co-receptor for HIV infection of CD4+ cells (Salazar *et al*., 2007; Sheffield *et al*., 2007).

1.2 *Treponema pallidum* subspecies *pallidum*

1.2.1 Taxonomy

The etiological agent of syphilis, *Treponema pallidum* subsp. *pallidum* is a spirochete bacterium that belongs to the family *Spirochaetaceae*, order *Spirochaetales,* class *Spirochaetia,* and phylum *Spirochaetes* (Whitman, 2010). As of the year more than 200 species or phylotypes have been identified, more than half of which are not yet cultured (Paster and Dewhirst, 2000). Members of the *Spirochaetes* phylum possess a unique cellular ultrastructure among bacteria with internal, periplasmic flagella (Whitman, 2010). Current analyses of 16S rRNA revealed four well delineated families within the order *Spirochaetales:* the families *Spirochaetaceae*, *Brevinemataceae*, *Brachyspiraceae,* and *Leptospiraceae* (Figure 3). The latter family contains *Leptospira interrogans* – the causative agent of leptospirosis. Family *Spirochaetaceae* comprises four genera: *Cristispira, Spirochaeta, Borrelia*, and *Treponema*. *Spirochaeta aurantia*

and *Spirochaeta thermophila* are two free-living representatives of the *Spirochaeta* genus (Canale-Parola and Breznak, 1975). Two clinically important members of the *Borrelia* genus include the causative agents of relapsing fever and Lyme disease, *Borrelia hermsii* and *Borrelia burgdorferi*, respectively (Schwan *et al*., 1989). The genus *Treponema* comprises three phylogenetic groups: the first one consists of *Treponema pallidum* (type species), T. *calligyrum,* T. *denticola,* T. *medium,* T. *phagedenis,* T. *putidum,* T. *refringens,* and T. *vincentii*; the second group consists of *Treponema amylovorum,* T. *berlinense,* T. *bryantii,* T. *brennaborense,* T. *lecithinolyticum,* T. *maltophilum,* T. *parvum,* T. *pectinovorum,* T. *porcinum,* T. *saccharophilum,* T. *socranskii*, and T. *succinifaciens*; the third group comprises *Treponema azotonutricium,* T. *primita,* T. *caldaria* and T. *stenostrepta* (Whitman, 2010). The Treponema genus, a member of which is the causative agent of syphilis, contains both free-living and host-associated members including both commensal and pathogenic species. Host-associated treponemes colonize various anatomical locations of insects and mammals including humans. *T. caldaria* is a fresh water thermophile (Abt *et al*., 2013), while *T. primitia* and *T. azotontricium* are commensals of the termite hindgut flora (Graber *et al*., 2004). *T. succinifaciens* are non-pathogenic inhabitants of the swine intestine (Han *et al*., 2011), while *T. pedis* and *T. brennaborense* are associated with bovine digital dermatitis (cattle lameness) (Evans *et al*., 2009). *T. viscentii, T. medium, T. denticola, T. lecithinolyticum, T. maltophilum,* and *T. socranskii* are associated with periodontal disease in humans (Dashper *et al*., 2011; Heuner *et al*., 2000; Lee *et al*., 2006; Park *et al*., 2002). *T. phagedenis,* once isolated from a human syphilis lesion and initially pathogenic to

rabbits, is now believed to be a non-pathogenic commensal of the genital flora of humans and other primates (Trott *et al*., 2003; Wallace and Harris, 1967).

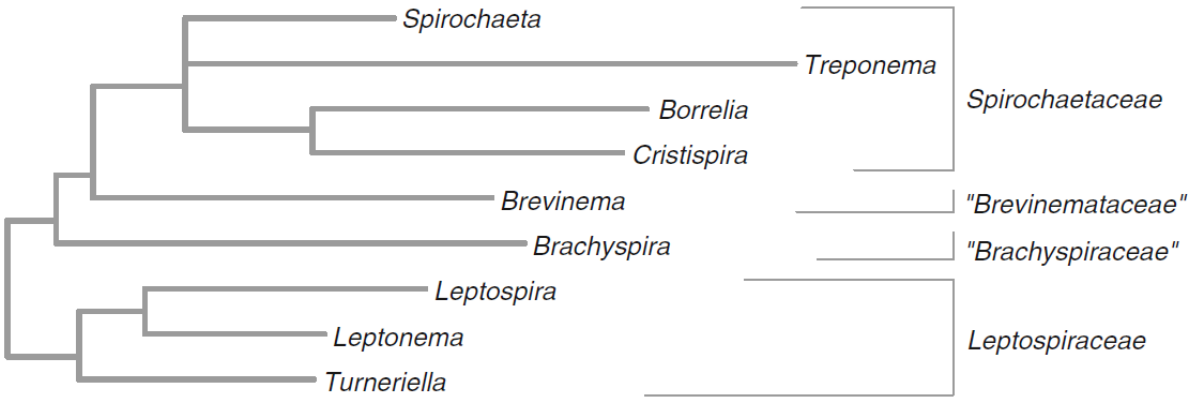


Figure 3. Relationships of Genera of Phylum *Spirochaetes* based upon 16S rRNA Analysis. (Whitman, 2010)

T. pallidum subspecies *pallidum* causes venereal syphilis, while the *T. pallidum* subspecies *pertenue* and *endemicum* cause the endemic, non-venereal diseases yaws and bejel (endemic syphilis), respectively (Antal *et al*., 2002). *Treponema carateum* causes another endemic human disease, pinta. *Treponema paraluiscuniculi*, a very closely related species, is the cause of venereal syphilis in rabbits, but is reportedly not infectious to humans (Centurion-Lara *et al*., 2013). These treponemes cannot be differentiated based on morphology and are very antigenically and genetically similar. Hence, the diseases caused by these treponemes are distinguished based on their epidemiological and clinical profiles (Antal *et al*., 2002). The *T. pallidum* subspecies cause chronic infections with early and late stages. However, they differ in the routes of transmission and the degree of dissemination capacity. The highly invasive *T. pallidum* subsp. *pallidum* can be transmitted via sexual contact, direct contact with primary or secondary lesions, or

vertically from a mother to her fetus, and the pathogen can disseminate to and infect any tissue or organ system, including the central nervous system. In contrast, *T. pallidum* subspecies *pertenue* and *endemicum* are transmitted via direct non-sexual contact and affect the skin, the mucous membranes, and bones with tissue destruction in late infection stages (Centurion-Lara *et al*., 2013). *T. carateum* causes severe skin pigmentation, but rarely involves tissue destruction. Unlike venereal syphilis, the latter three infections are thought to not involve the central nervous system or the fetus (Antal *et al*., 2002).

1.2.2 Biology of *T. pallidum*

Treponema pallidum subspecies *pallidum* (hereafter referred to as *Treponema pallidum*) is a helically shaped bacterium that is 6 to 15 µm in length and 0.2 µm in diameter (Lafond and Lukehart, 2006). Since these bacteria are very thin compared to *Escherichia coli*, which are 0.5 µmwide (Kubitschek, 1990), darkfield microscopy is used to visualize *T. pallidum* for clinical and research purposes. The microorganism is surrounded by a cytoplasmic membrane, and further enclosed by a loosely attached outer membrane (Lafond and Lukehart, 2006). A thin layer of peptidoglycan is found between the two membranes. *T. pallidum* is a highly motile microorganism that propels itself via a corkscrew-like mechanism by rotations around its longitudinal axis. The particular mode of motility is due to the use of endoflagella, organelles located in the periplasmic space between the inner and outer membranes of *T. pallidum* (Izard *et al*., 2009; Liu *et al*., 2010). The helical structure, internal flagella, and corkscrew-like motility of the pathogen allow enhanced motility through viscous mucous and tissue membranes found in the host

(Lafond and Lukehart, 2006). An electron microscopic image of *T. pallidum* is presented in Figure 4.

Figure 4. Electron microscopic image of *Treponema pallidum* in a human primary lesion (Drusin *et al*., 1969)

1.2.2.1 Genome sequencing: limited metabolic capacity

The complete genome sequencing of *T. pallidum* confirmed a rather small genome of 1.14 Mb that encodes 1,041 putative proteins (Fraser, 1998). Upon comparison of genome sizes, very few bacteria have genomes smaller than that of *T. pallidum* (Lafond and Lukehart, 2006). For example, the genome of a typical gram-negative bacterium, such as *Escherichia coli* K-12, is 4.6 Mb and that of a conventional gram-positive bacterium, such as *Bacillus subtilis*, is 4.2 Mb. A striking lack of metabolic capabilities was also discovered by the sequencing of the *T. pallidum* genome (Fraser, 1998). While *T. pallidum* is capable of carrying out glycolysis, analysis of the genome sequencing results indicated a lack of tricarboxylic acid cycle enzymes and components

of the electron transport chain. Absence of pathways for the synthesis of enzyme cofactors and nucleotides, as well as amino acid and fatty acid synthesis, further confirmed the scarcity of biosynthetic pathways and suggested that *T. pallidum* relies on the host to derive most of the essential macromolecules (Fraser, 1998; Lafond and Lukehart, 2006).

1.2.2.2. *In vivo* and *in vitro* growth of *T. pallidum*

A number of factors have hindered investigation of the biology and biochemistry of *T. pallidum*, among which is its slow generation time (Lafond and Lukehart, 2006). The *in vitro* generation time of *T. pallidum* was found to be 30 to 50 hours (Cumberland and Turner, 1949) and inoculation experiments in rabbits demonstrated that the pathogen divides every 30-33 hours (Cumberland and Turner, 1949; Fieldsteel *et al*., 1981). Perhaps the greatest impediment to studies of the microorganism is the inability of *T. pallidum* to survive and multiply outside of its mammalian host (Ho and Lukehart, 2011; Lafond and Lukehart, 2006). Propagation of *T. pallidum* to more than 100-fold in tissue culture has not been successful, which equates to about seven generations (Fieldsteel *et al*., 1981, 1982; Norris, 1982). The pathogen loses its infectious capability soon after harvest within hours or days. Since culturing *in vitro* is only possible transiently in rabbit epithelial cells, *T. pallidum* must be propagated in rabbits for laboratory studies (Ho and Lukehart, 2011). In addition, *T. pallidum* lacks enzymes, such as catalase and oxidase, for detoxification of reactive oxygen species and is thus sensitive to oxygen exposure. Furthermore, the pathogen is sensitive to temperatures higher than 37°C since the typical

σ32-regulated heat shock response is lacking in *T. pallidum* (Fieldsteel *et al*., 1981; Fraser, 1998; Haake and Lovett, 1994a).

1.2.2.3 *Treponema pallidum* cell envelope ultrastructure

Although, as noted earlier, *T. pallidum* shares structural similarities to typical Gram-negative bacteria, such as having outer and inner membranes and a periplasmic space, it lacks lipopolysaccharide (LPS) (Bailey *et al*., 1985; Radolf and Norgard, 1988). In contrast to traditional Gram-negative bacteria, the peptidoglycan layer in *T. pallidum* was originally believed to directly overlie the cytoplasmic membrane (Holt, 1978). As a result, a large periplasmic space was created which contained the internal flagella, or "axial filaments" wrapped around the cytoplasmic body, responsible for the characteristic corkscrew motility (Jepsen *et al*., 1968; Johnson *et al*., 1973; Swain, 1955). Recently, cryo-electron tomography studies of the outer envelope of *T. pallidum* showed a thin layer of peptidoglycan between the membranes that ensures structural stability while permitting flexibility (Izard *et al*., 2009; Liu *et al*., 2010). The peptidoglycan layer in *T. pallidum* divides the periplasmic space into two distinct regions. The space between the peptidoglycan and the outer membrane contains the flagellar filaments that originate as flagellar bundles at flagellar motors on each end of the bacterium, wind around the flexible protoplasmic cell cylinder, and overlap in the middle. The flagellar filaments in *T. pallidum* are composed of several major proteins, such as the sheath protein FlaA and flagellar core proteins FlaB1, FlaB2 and FlaB3 (Liu et al., 2010). The region between the inner membrane and the peptidoglycan layer contains abundant lipoproteins anchored in

the outer leaflet of the inner membrane. In agreement with previous studies, patches of density were observed along the inner leaflet of the outer membrane as well, suggesting the presence of lipoproteins anchored in the inner leaflet of the outer membrane with their protein moieties in the periplasmic space (Hazlett *et al*., 2005; Liu *et al*., 2010). A model of the molecular architecture of the *T. pallidum* cell envelope based on cryo-electron tomography studies is shown in Figure 5.

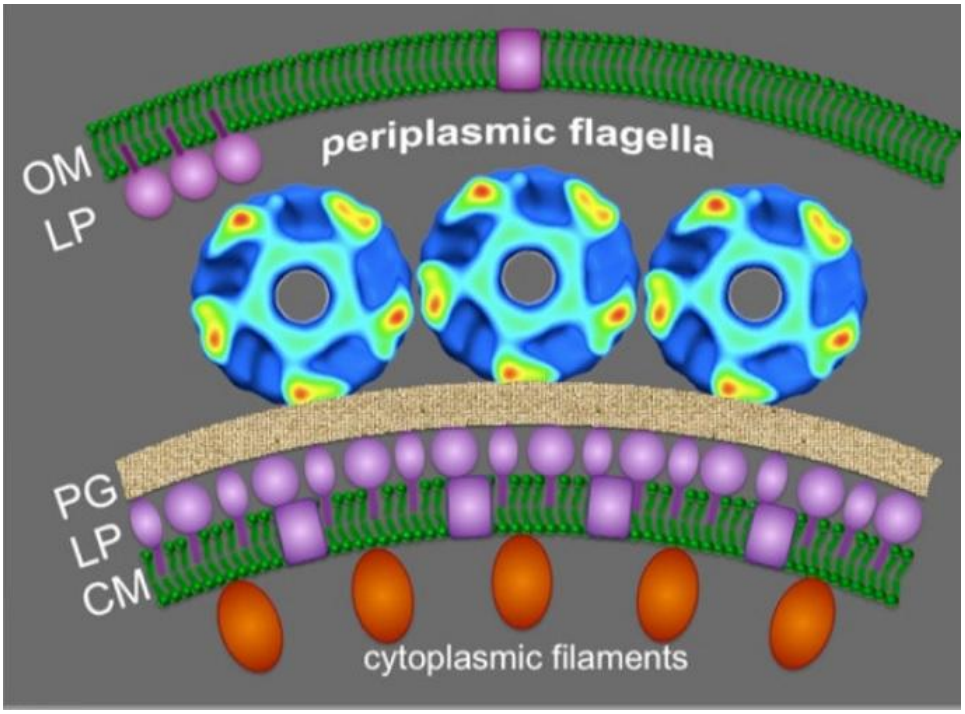


Figure 5. Model of the *T. pallidum* cell envelope architecture constructed from cry-electron tomography imaging of *T. pallidum*. OM: outer membrane, LP: lipoprotein, PG: peptidoglycan, CM: cell (inner) membrane. (Liu *et al*., 2010)

The absence of LPS and the loosely associated outer membrane make *T. pallidum* much more fragile than typical Gram-negative bacteria. Use of low concentrations of detergents and physical manipulations such as centrifugation, resuspension, and washing can easily disrupt the outer membrane of *T. pallidum* (Bailey *et al*., 1985; Radolf and Norgard, 1988). The inability to culture the organism *in vitro* coupled with the fragility

of its outer membrane make *T. pallidum* genetically and biochemically intractable – a major hindrance to the molecular research of the pathogen (Lafond and Lukehart, 2006).

1.2.2.4 *Treponema pallidum* outer membrane proteins

Surface exposed outer membrane proteins (OMP) are often involved in interactions with the host, such as during the process of pathogenesis. Examples include the OMP VacA, an extracellular cytotoxin produced by50-60% of *Helicobacter pylori,* and the OMP OmpA, an outer membrane protein that contributes to serum resistance in *E. coli,* both of which proved essential for the pathogenesis of these organisms (Keenan *et al*., 2000; Weiser and Gotschlich, 1991). The surface localization of outer membrane proteins is of great medical importance as well, since they often represent good vaccine candidates. Successful use of OMPs for vaccine production was demonstrated with *Burkholderia multivorans* OMPs, for instance, in providing protection against subsequent infections with *B. multivorans* (Bertot *et al*., 2007).

Early researchers demonstrated that intact *T. pallidum* was poorly antigenic as antibodies from the sera of infected animals did not readily bind to the treponemes (Deacon *et al*., 1957). Unlike traditional Gram-negative bacteria, *T. pallidum* lacks lipopolysaccharide (Bailey *et al*., 1985); in addition, early freeze-fracture and freeze-etch electron microscopy studies have demonstrated that the outer membrane of *T. pallidum* contains a paucity of integral outer membrane proteins (Cox *et al*., 1992; Radolf *et al*., 1989; Walker *et al*., 1989). In later experiments, outer membrane vesicles isolated from *T. pallidum* were analyzed for integral proteins by freeze fracture electron microscopy and tested for antigens by immunoblot analysis (Blanco *et al*., 1994; Radolf *et al*., 1995).

The latter, as well as more recent scanning probe microscopy studies of the outer surface of the treponemes (Liu *et al*., 2010), confirmed the paucity of integral OMPs in *T. pallidum*. The characteristic poor antigenicity of the pathogen, which earned *T. pallidum* the name "the stealth pathogen" (Salazar *et al*., 2002), may aid the organism in evading immune detection while establishing a persistent disseminated infection.

Earlier freeze-fracture analyses revealed that the number of integral outer membrane proteins of both *T. pallidum* subsp. *pallidum* and *pertenue* were roughly a hundred-fold lower than that of *E. coli* (Radolf *et al*., 1989; Walker *et al*., 1991), twenty-fold lower than that of *B. burgdorferi*, the causative agent of Lyme disease (Walker *et al*., 1989, 1991), and ten-fold lower than that of *T. phagedenis*, a non-pathogenic human-associated treponeme originally isolated from a syphilis sore (Walker *et al*., 1989; Wallace and Harris, 1967). Interestingly, the numbers were similar between *E. coli* and *S. aurantia*, a free-living spirochete. Taken together these observations suggest that the low integral outer membrane protein density of *T. pallidum* is not a structural requirement within spirochetes, but instead may be an evolutionary adaptation for host immune evasion (Walker *et al*., 1989). Figure 6 compares the freeze fracture faces of the *E. coli* and *T. pallidum* outer membranes and the integral membrane protein content in each of them.

The rare *T. pallidum* OMPs are likely very important in the pathogenicity of the organism and in the interaction with the host immune system; therefore they may constitute an effective syphilis vaccine. For these reasons, identification and characterization of these rare outer membrane proteins and their role in *T. pallidum*'s

pathogenicity have been the subject of intense research over the last two decades (Cameron and Lukehart, 2013; Lafond and Lukehart, 2006; Radolf, 1995)

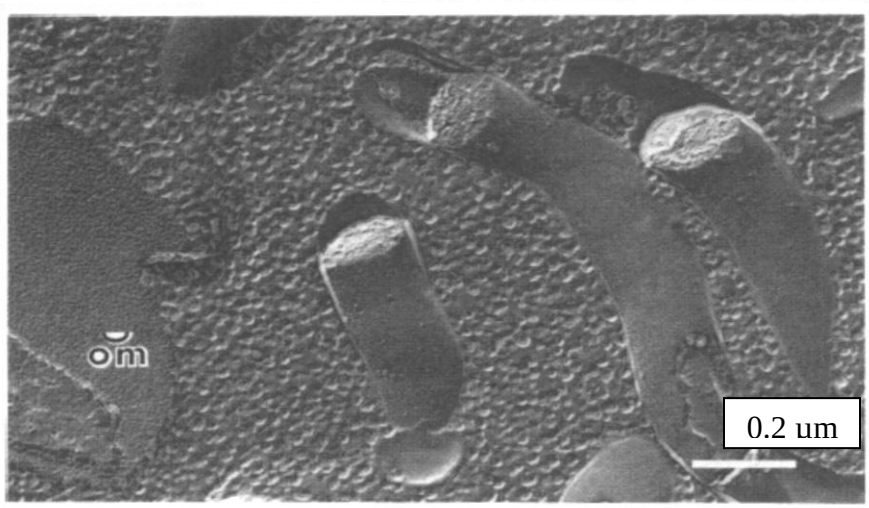


Figure 6. Outer membrane freeze fractures of *Escherichia coli* (left) and *Treponema pallidum(right).* The outer leaflet of the outer membrane (om) of *E. coli* shows a uniformly dense distribution of outer membrane proteins in sharp contrast to the scarce particles in the fracture faces of the *T. pallidum* outer membrane (Radolf *et al*., 1989).

1.2.2.5 Identification of *Treponema pallidum* outer membrane proteins

In the past, various techniques were used in an attempt to identify the rare *T. pallidum* outer membrane proteins. Early studies used phase partitioning with different detergents (Bailey et al., 1985; Cunningham et al., 1988; Penn et al., 1985), acid-mediated separation of membranes (Stamm and Bassford, 1985), or density gradient ultracentrifugation of organisms lysed in a hypotonic solution (Alderete and Baseman, 1980). Although a number of proteins were initially believed to be surface-exposed based on these methods, it was later demonstrated by Cox *et al.* that physical

manipulations such as centrifugation, resuspension, and non-ionic detergents damaged the *T. pallidum* outer membrane exposing proteins that were not originally found on the surface (Cox *et al*., 1992).

Since the completion of the first whole genome sequencing of *T. pallidum* (Fraser, 1998), researchers have been utilizing bioinformatics to identify potential surface exposed proteins. Analyses of protein sequences for predicted cleavable signal sequences and transmembrane domains identified several proteins as candidate outer membrane proteins (Cameron, 2003; Centurion-Lara *et al*., 1999). Three of the proteins predicted by the bioinformatics analyses to be surface exposed, Tp0155, Tp0483, and Tp0751 (pallilysin), were recombinantly expressed and shown to bind to host extracellular and coagulation cascade components, such as fibronectin (Tp0155 and Tp0483), laminin and fibrinogen (pallilysin) (Cameron *et al*., 2004, 2005; Houston *et al*., 2011). Additional studies using synthetic peptides systematically localized the minimum laminin-binding region of pallilysin to 10 amino acids: amino acids P^{98}, V^{99}, Q^{100}, T^{101}, amino acids W^{127} and I^{128}, and amino acids T^{182}, A^{183}, I^{184}, and S^{185} (Cameron et al., 2005). Pallilysin was later confirmed by opsonophagocytic assays to be surface-exposed (Houston *et al*., 2012). A novel proteolytic function was discovered in addition to its role as an adhesin, making it the first known *T. pallidum* protease capable of degrading host components (Houston *et al*., 2011). Another putative *T. pallidum* OMP, Tp0136, was also identified as a fibronectin binding adhesin (Brinkman *et al*., 2008). Therefore, the binding of various host components by multiple adhesins on the surface of *T. pallidum* suggests they play an essential role in the organism's pathogenesis. Another predicted outer membrane protein, Tp0897 (TprK), was demonstrated to be a target of opsonic antibodies and was expressed

preferentially in the course of infection (Centurion-Lara *et al*., 1999). In addition, TprK undergoes antigenic variation as a means of evading the host adaptive immune response (Centurion-Lara *et al*., 2013).

1.3 Dissemination Capacity of *T. pallidum*

The syphilis treponeme is considered one of the most invasive pathogens known and is able to invade and persist in a wide variety of tissues and organs (Lafond and Lukehart, 2006).The diverse disseminated clinical manifestations of secondary, tertiary, and congenital syphilis in humans demonstrate the invasiveness of *T. pallidum*. While infection initially occurs at ano-genital or, more rarely, oral and non-genital dermal sites, the generalized rash of secondary syphilis clearly demonstrates that organisms disseminate widely from the primary site of infection (Lafond and Lukehart, 2006). Direct detection of *T. pallidum* in tissues and fluids distal from the initial site of infection provides further support for the highly invasive capabilities of the pathogen. PCR analyses and infectivity testing have shown that *T. pallidum* is found in the cerebrospinal fluid (CSF) of individuals with early and latent syphilis (Lukehart *et al*., 1988; Marra *et al*., 2004). The organism has been detected in tertiary dermal gummatous lesions decades after initial infection by silver staining and immunofluorescence microscopy (Handsfield *et al*., 1983; Kampmeier, 1964; Lafond and Lukehart, 2006) and by PCR (Zoechling *et al*., 1997).

Direct evidence for the ability of *T. pallidum* to invade a wide variety of different tissue types is also provided by numerous studies of experimental infections of model animals. After intradermal infection of rabbits, the pathogen was demonstrated by

microscopy to have disseminated to the skin, testes, spleen, and lymph nodes (Sell *et al*., 1980). Following intratesticular infection of rabbits, early studies detected skin and bone lesions (Turner, 1957) and treponemes were found in the lymph nodes, brain, and aqueous humor, and in the cerebrospinal fluid (CSF) as early as 18 hours after inoculation (Collart *et al*., 1971). After intravenous infection of rabbits, *T. pallidum* RNA was detected in the CSF by reverse transcription-PCR (Tantalo *et al*., 2005). Intrathecal inoculation resulted in ocular syphilis in 6% of rabbits (Marra *et al*., 1991), demonstrating that *T. pallidum* can travel from the CSF to the eye.

Furthermore, dissemination of the pathogen is not only systemic, but also rapid. Rabbit inoculation studies showed that *T. pallidum* enters the bloodstream within minutes of intratesticular or intradermal inoculation (Cumberland and Turner, 1949; Stokes *et al*., 1944), and organisms applied to mucosa are found in deeper tissues within hours of infection (Mahoney and Bryant, 1934). Finally, blood from mice, monkeys, and rabbits during early stages of infection proved to be infectious to naive animals, which is an indication that *T. pallidum* is found in the bloodstream of the infected host facilitating the establishment of a widespread disseminated infection.

The first step in *T. pallidum* invasion and dissemination, however, is the attachment of the treponemes to host cells. Attachment of the pathogens to the vasculature facilitates intra- and extravasaion and subsequent penetration of tissues to establish a disseminated infection. Adhesion studies demonstrated that *T. pallidum* is able to attach to a wide variety of cell types including epithelial, fibroblast-like, and endothelial cells of both rabbits and humans (Fitzgerald *et al*., 1977a; Hayes *et al*., 1977; Lee *et al*., 2003; Lovett *et al*., 1988). Moreover, *ex vivo* studies showed the ability of *T.*

pallidum to adhere to isolated capillaries (Quist *et al*., 1983), and abdominal walls (Riviere *et al*., 1989). Attachment to host cellular components has been demonstrated for other pathogenic spirochetes as well, including *T. denticola* (Peters *et al*., 1999), *B. burgdorferi* (Comstock and Thomas, 1989; Grab *et al*., 2009), and *Leptospira* species (Thomas and Higbie, 1990). In contrast, the non-pathogenic treponeme *T. phagedenis* did not adhere to cultured cells (Fitzgerald *et al*., 1977a, 1977b; Hayes *et al*., 1977), an indication that attachment is characteristic of, and specific to, pathogenic treponemes. Further support was gained from experiments in which heat-killed or non-motile *T. pallidum* were unable to attach to cells (Fitzgerald *et al*., 1975, 1977a). Interestingly, the fact that immune sera from either humans or rabbits interfered with the adherence of the viable pathogens to cell cultures, suggested that attachment to host cells may be mediated by specific *T. pallidum* surface-exposed antigens (Fitzgerald *et al*., 1977a; van der Sluis *et al*., 1987). In addition to attachment to cell cultures, *T. pallidum* was demonstrated to bind to a variety of host serum and cell components, as well as the extracellular matrix (ECM) including fibronectin, laminin, fibrinogen, collagen I, and hyaluronic acid (Cameron, 2003; Cameron *et al*., 2004; Fitzgerald and Repesh, 1985; Fitzgerald *et al*., 1984; Houston *et al*., 2011). Adherence to ECM components has been shown to mediate the attachment of other pathogenic bacteria to host cells (Finlay and Falkow, 1997). The major sheath protein of the related oral spirochete *T. denticola*, for example, adheres to fibronectin and laminin (Fenno *et al*., 1996). A wide variety of adhesins mediate attachment of *Streptococcus pyogenes* to host ligands that may also allow the pathogen to penetrate a wide variety of tissue types (Patti *et al*., 1994). Similarly, *T. pallidum*'s adhesins that bind to different ECM components may facilitate the overwhelming

capacity of the organism to penetrate a variety of different tissues and establish a widespread disseminated infection (Ho and Lukehart, 2011).

In addition, *T. pallidum* has been shown to induce the production of active matrix metalloproteinase-1 (MMP-1) in dermal cells (Chung *et al*., 2002). The fact that MMP-1 is involved in the degradation of collagen type I, the most abundant component of the human dermis, suggests that activation by *T. pallidum* may assist in tissue penetration. Moreover, the characteristic corkscrew motility of the pathogen allows it to move easily in gel-like materials, such as the connective tissue, and may further enhance its dissemination capabilities. Along these lines, motility has been established as a virulence factor for a number of other bacterial pathogens (Josenhans and Suerbaum, 2002; Lux *et al*., 2001).

Following initial infection, *T. pallidum* quickly gains access to deeper tissues and the bloodstream (Mahoney and Bryant, 1934; Stokes *et al*., 1944), which implies a hematogenous route of dissemination. Hematogenous spread, in turn, requires that the treponemes traverse the endothelial barrier from the vascular lumen to the surrounding tissues (Haake and Lovett, 1994b). To breach the endothelial barrier an extracellular pathogen must either use the trans- or paracellular pathway (Nassif *et al*., 2002). The former requires transcytosis, while the latter involves opening the intercellular junctions that can be impermeable even to small molecules including water. Previous studies showed that *T. pallidum* was able to traverse the junctions between cultured endothelial cells (Lovett *et al*., 1988; Riley *et al*., 1992). Electron microscopy images of primary and secondary syphilis skin lesions showed that *T. pallidum* organisms were located mainly in the blood vessel walls and dermal tissue of the chancre lesions, but also suggested that

the pathogen may use transcytosis to spread through the endothelium (Juanpere-Rodero *et al*., 2013).

1.3.1 Endothelial barrier

The endothelium, which lines the luminal side of all vessel types, is comprised of a monolayer of vascular endothelial cells with cobblestone morphology (Huber, 2009) that regulate the flow of nutrients, various biologically active molecules, and blood cells (Cines *et al*., 2014). The average human endothelial surface area can be assumed to exceed 1000 m^2, which is approximately 600-times that of the epidermis (Müller and Griesmacher, 2000). In the course of a blood-borne infection, this significant surface is exposed to a large number of microorganisms. The traditional view deemed endothelial cells a passive lining of the vasculature that served to contain blood cells and plasma (Mantovani *et al*., 1992). In contrast, more recent studies showed that upon exposure to environmental stimuli, such as microbial components, endothelial cells undergo profound changes in gene expression that result in the active participation of endothelial cells in immunity, inflammation, and thrombosis (Henneke and Golenbock, 2002). For instance, on binding of microbial substructures, such as endotoxin from Gram-negative bacteria, endothelial cells elicit a signal that results in the formation of cytokines and the expression of cellular adhesion and pro-coagulant molecules. Inflammatory and immune cells then migrate from the bloodstream to the site of infection (Lafond and Lukehart, 2006). Despite the lack of lipopolysaccharide, virulent *T. pallidum* was shown to induce expression in cultured endothelial cells of the adhesion molecules ICAM-1 (Intercellular adhesion molecule 1) and E-selectin (169, 255). While these were also activated by the *T.*

pallidum lipoprotein TpN47, no activation occurred by heat-killed *T. pallidum* or the non-pathogenic treponeme *T. phagedenis* (169, 255). Therefore, endothelial cell activation is likely a pathogen-specific, active process mediated by specific *T. pallidum* molecules (Lafond and Lukehart, 2006).

Endothelial cell are connected through numerous transmembrane adhesive proteins that are either found in junctional structures or along the intercellular cleft (Huber, 2009). These proteins provide cell to cell adhesion and control vascular permeability to fluids, molecules, and transmigrating leukocytes. The endothelial permeability can be influenced by specific permeability increasing agents, such as histamine or thrombin, or by inflammatory cytokines that can cause gaps to form at the intercellular contacts (Dejana and Orsenigo, 2013). Adherens junctions and tight junctions are among the best-studied junctional complexes in the endothelium. Adherens junctions are formed by cadherins, which are cell adhesion molecules that mainly support binding between similar molecules on opposing cells (Vestweber, 2000). In contrast to epithelial cells, endothelial cells do not have classic desmosomes; instead, they express the desmosomal protein desmoplakin (Vestweber, 2000). Tight junctions serve to seal the paracellular space and, unlike in epithelia, can be found intermingled with adherens junctions in endothelia (Anderson and Van, 1995). Multiple integral membrane proteins have been identified to comprise tight junctions: occludin, claudins and junction-associated membrane (JAM) proteins (Martìn-Padura *et al*., 1998; Nassif *et al*., 2002). From a structural perspective, tight junctions form an uninterrupted network of parallel, interconnected, intramembrane strands of protein resulting in a series of multiple barriers (Schneeberger *et al*., 1978). Different tissues have tight junctions that exhibit varying

degrees of cytoplasmic fibril organization as measured by transendothelial electrical resistance (TEER). A correlation exists between increased cytoplasmic fibril organization (high TEER) and decreased membrane permeability (Claude, 1978). The difference in the electrical resistance values in endothelia depending on location implies important functional consequences. For instance, human placental endothelial cells exhibit TEER values of 22–52 OhmXcm^2 (Jinga *et al*., 2000), which is in accordance with the fast paracellular exchange of nutrients and waste between the mother and fetus. The blood-brain barrier has a much higher resistance (1500–2000 OhmXcm^2) to paracellular diffusion, which serves to maintain the tightly-regulated brain homeostasis (Huber *et al*., 2001).

1.3.1.1 Basement Membrane and Endothelial Cell Behaviour

Basement membranes are extracellular substrata that line the basal surface of endothelial cells throughout the entire vascular system (Grant *et al*., 1990). In addition, these matrices are also closely associated with epithelia, smooth and skeletal muscle, and the nervous system. The molecular composition of basement membranes consists of specific and constant components, such as collagen IV, laminin, entactin, fibronectin, and heparan sulfate (Grant *et al*., 1990). Basement membranes form a sleeve around the endothelium of capillaries, arterioles and venules and support the vascular architecture, maintain cell polarity of the vessel and regulate endothelial cell behaviour, such as cell adhesion, differentiation and proliferation (Dejana *et al*., 1988; Folkman and Klagsbrun, 1987; Madri *et al*., 1988).

When endothelial cells are cultured on tissue culture-treated plastic, they exhibit characteristic cobblestone morphology and are contact-inhibited (Grant *et al*., 1990). When cultured on other substrata, such as collagen I or fibronectin, an increase in proliferation is observed, while a laminin substratum promotes endothelial differentiation into capillary-like structures (Grant *et al*., 1989; Kubota *et al*., 1988). Capillary tube formation is also observed when endothelial cells are grown to post-confluent state, in the presence of tumor-conditioned media, or after a period of 4 to 5 weeks without endothelial cell growth factors (Folkman and Klagsbrun, 1987; Grant *et al*., 1990; Maciag *et al*., 1982). When human umbilical vein endothelial cells (HUVEC) are cultured on Matrigel, a reconstituted basement membrane matrix, tube formation occurs within 18 hours (Grant *et al*., 1990). Varying the level of thickness of the Matrigel can influence the morphology of the endothelial cells cultured upon it, with a thin layer promoting attachment and proliferation, rather than differentiation into capillary tubes (BD Biosciences, 2011).

1.4 Pallilysin, a *Treponema pallidum* Surface-Exposed Adhesin and Protease

Dissemination of *T. pallidum* to establish a systemic infection requires the pathogen to first traverse the endothelial layer and the basement membrane lining the vasculature. Indeed, the treponemes have a predilection for perivascular areas during infection and were shown to attach to isolated rabbit capillaries (Quist *et al*., 1983). *T. pallidum* also specifically interacts with cultured vascular endothelial cell layers (Lee *et al*., 2003). Previous studies have demonstrated attachment to and penetration of cultured endothelial monolayers through their intercellular junctions (Lovett *et al*., 1988; Thomas

et al., 1989). In addition, Fitzgerald *et al*. (1984) demonstrated that *T. pallidum* also binds to the basement membrane underlining the vascular endothelial cell. Specific attachment to the structural components of basement membranes and various extracellular matrix components, such as fibronectin, laminin, collagen, IV, collagen I, and hyaluronic acid, were reported suggesting the involvement of specific outer membrane adhesins in ECM binding and dissemination (Cameron, 2003; Fitzgerald and Repesh, 1985; Fitzgerald *et al*., 1984; Ho and Lukehart, 2011; Lee *et al*., 2003).

Despite the paucity of OMPs on the surface of *T. pallidum* (Liu *et al*., 2010; Radolf *et al*., 1989; Walker *et al*., 1989), bioinformatics analyses of the *T. pallidum* genome for OMPs and subsequent ECM attachment assays identified, among a few others, a specific laminin-binding adhesin, pallilysin (Cameron, 2003). The *T. pallidum* adhesin, pallilysin, was found to be expressed during infection and showed strong affinity for laminin, the major glycoprotein found within basement membranes (Cameron, 2003). Pallilysin was shown to bind to a wide variety of laminin isoforms, consistent with a highly invasive pathogen that penetrates several basement membranes throughout the course of infection, such as during intravasation, extravasation, and invasion of other tissues (Cameron *et al*., 2005). Further support for the role of the *T. pallidum* adhesin in dissemination of the pathogen via the circulatory system was given by its ability to also bind specifically to fibrinogen, a key structural protein in blood coagulation (Houston *et al*., 2011).

Analysis of the pallilysin primary sequence revealed a C-terminal putative HEXXH metalloprotease motif, which is found in approximately 50% of known metalloendopeptidases (Houston *et al*., 2011). Bioinformatics analyses also revealed that,

however, only 3% of all *T. pallidum* proteins contain that amino acid sequence (Houston, personal communication, 2014). *In vitro* degradation assays confirmed that recombinant pallilysin degrades both human fibrinogen and laminin; thus, pallilysin presented the first *T. pallidum* protease capable of degrading host components. Pallilysin was shown to bind zinc and was inhibited by metalloprotease inhibitors (1,10-phenanthroline), but not by serine or cysteine protease inhibitors. In a later study, site-directed mutagenesis of the pallilysin HEXXH motif (AEXXH [H198A], HAXXH [E199A], and HEXXA [H202A]) abolished pallilysin-mediated fibrinogenolysis without a negative effect on host component binding confirming that the HEXXH motif was part of the active site of the protease (Houston *et al*., 2012). Furthermore, the dual role of pallilysin as an ECM binding adhesin and a zinc metalloprotease provided a novel paradigm in the pathogenesis of *T. pallidum* and the establishment of a disseminated infection.

In addition to degrading soluble fibrinogen, wild-type pallilysin was also able to degrade insoluble fibrin clots. Thrombin, a component of the coagulation cascade, converts soluble fibrinogen to insoluble fibrin, a major constituent of blood clots, which serves to contain bacteria. Several other pathogenic bacteria, including Group B streptococcus and *Staphylococcus aureus*, have developed proteolytic solutions to overcome this host defence mechanism (Harris *et al*., 2003; Ohbayashi *et al*., 2011).

Wildtype pallilysin, but not the HEXXH mutants, was further shown to undergo autocatalytic cleavage of its N-terminal pro-domain in order to achieve full proteolytic activity (Houston *et al*., 2012). N-terminal sequencing was utilized to identify the pallilysin autocatalytic cleavage sites, T^{46}-A^{47} and Q^{92}-T^{93}, resulting in 26 and 18 kDa wild-type pallilysin proteins generated from the 32 kDa (237 amino acid) full-length

pallilysin. Pre-incubation of the full-length pallilysin for 21 hours generated the final autocatalytically cleaved mature protease (18kDa pallilysin, T^{93}-P^{237}) resulting in more rapid fibrinogenolysis. Interestingly, *in vitro* studies showed that thrombin was also able to cleave pallilysin generating a truncated fragment with enhanced proteolytic activity. Therefore, while pallilysin can activate itself through inter-molecular autocatalysis, it could also hijack the host coagulation cascade to facilitate protease activation.

Finally, using opsonophagocytosis assays on viable *T. pallidum*, pallilysin was demonstrated to be target of opsonic antibodies confirming its surface-exposed cellular localization as predicted by the initial bioinformatics analysis of its primary sequence. In summary, pallilysin is a *T. pallidum* surface-exposed adhesin and a metalloprotease with an HEXXH active site motif that requires autocatalytic or host-mediated cleavage to achieve full proteolytic activity.

1.4.1 Heterologous expression of pallilysin in *T. phagedenis*

Research on the pathogenic mechanisms used by *T. pallidum* to establish a persistent disseminated infection has been hindered by the fact that the organism cannot be continuously cultured *in vitro* (Fieldsteel *et al*., 1981, 1982; Norris, 1982). Since *T. pallidum* is not amenable to genetic manipulations by traditional experimental methods, direct investigation of the function of individual gene products important for pathogenesis remains impossible. As an alternative, heterologous expression of *T. pallidum* genes has been accomplished in *E. coli* (Hansen *et al*., 1985; Swancutt *et al*., 1989) and *T. denticola* (Chi *et al*., 1999; Slivienski-gebhardt *et al*., 2004); however, neither were particularly suited for the investigation of virulence factors associated with

T. pallidum pathogenesis. The dissimilarity in the outer membrane ultrastructure and physiology between *T. pallidum* and *E. coli* resulted in limited functional data obtained (Isaacs and Radolf, 1990). Although *T. denticola* is a cultivable treponeme and thus heterologous expression studies are more relevant, it is a human-associated pathogenic treponeme found in subgingival plaques associated with periodontitis (Chi *et al*., 1999; Lux *et al*., 2001). The fact that *T. denticola* attaches to host cells and cellular components (Haapasalo *et al*., 1991; Olsen, 1984; Peters *et al*., 1999), penetrates endothelial cell monolayers (Lux *et al*., 2001; Peters *et al*., 1999), and invades the gingival connective tissue (Frank, 1980; Peters *et al*., 1999; Riviere *et al*., 1991) complicates investigations of heterologously expressed virulence factors predicted to be involved in *T. pallidum* adhesion, tissue invasion, and dissemination.

In contrast, *T. phagedenis* is a human-associated non-pathogenic culturable member of the *Treponema* genus. A strict anaerobe that does not attach to or invade cell monolayers (Fitzgerald *et al*., 1977a; Hayes *et al*., 1977; Lovett *et al*., 1988; Peters *et al*., 1999), *T. phagedenis* was a suitable candidate for heterologous expression of *T. pallidum* virulence factors associated with dissemination of the pathogen (Cameron et al., 2008a). With a similar GC ratio to that of *T. denticola*, *T. phagedenis* could successfully be transformed with a shuttle vector developed for use in *T. denticola*. The *T. pallidum* flagellar gene *flaA* was previously expressed in *T. denticola* using the *E. coli*-*T. denticola* shuttle vector pKMR4PE, henceforth referred to as pKMR (Chi *et al*., 1999). The entire pallilysin ORF, including the putative signal sequence and upstream ribosome-binding site, were thus cloned into the multiple cloning site of pKMR downstream of an erythromycin resistance (emr) gene cassette (Cameron et al., 2008a). Together, the

pallilysin and the emr gene cassette were located downstream of the constitutively expressed *T. denticola* protease, prtB, promoter (Arakawa and Kuramitsu, 1994).

Heterologous expression of pallilysin on the surface of *T. phagedenis* was confirmed via immunofluorescence analysis with pallilysin-specific antibodies (Cameron *et al*, 2008). Laminin attachment assays demonstrated that heterologous expression of pallilysin conferred upon *T. phagedenis* the capacity to attach to laminin. Houston and coworkers showed in a later study that *T. phagedenis* heterologously expressing pallilysin were able to degrade fibrin clots, similar to experiments with recombinant pallilysin. Moreover, click-chemistry-based palmitoylation profiling of heterologously expressed pallilysin showed that pallilysin was S-palmitoylated in *T. phagedenis* (Houston *et al*., 2012). Although recent cryo-electron tomography studies of the outer envelope of *T. pallidum* failed to detect surface-exposed outer membrane lipoproteins (Liu *et al*., 2010), heterologous expression of pallilysin (Cameron et al., 2008a; Houston et al., 2011), together with the more recent opsonophagocytic assays (Houston *et al*., 2012), confirmed that pallilysin is a surface-exposed lipoprotein. These results coupled with the fact that pallilysin is expressed during infection (Cameron, 2003) and is capable of both binding to and degrading host ECM components (Cameron *et al*., 2005; Houston *et al*., 2011, 2012) strongly advocate a role for pallilysin in the dissemination of *T. pallidum*.

1.5 Research Hypothesis and Objective

Treponema pallidum is one of the most invasive pathogens known, being able to spread via the bloodstream to every organ system in the human host and cross the placental and blood-brain barriers (Lafond and Lukehart, 2006). Hematogenous spread

requires, however, that the microorganism cross the endothelial barrier of the vasculature and the underlining basement membrane (Haake and Lovett, 1994a). Indeed, *T. pallidum* has been shown to attach to isolated capillaries, cultured vascular endothelial cells, basement membranes and components of the extracellular matrix, such as laminin and fibronecting among others (Cameron *et al*., 2004, 2005; Fitzgerald *et al*., 1984; Quist *et al*., 1983). In addition, the pathogenic treponemes were shown to be able to cross endothelial monolayers most likely through the intercellular junctions (Thomas *et al*., 1989). Specific binding is likely mediated by surface exposed outer membrane adhesins that allow *T. pallidum* to attach to the vasculature facilitating intravasation and extravasation and penetration of other tissues to establish a widespread disseminated infection. Several predicted rare outer membrane proteins in *T. pallidum* have been shown to specifically bind to extracellular matrix components; one of them, pallilysin, was demonstrated to be surface exposed in *T. pallidum* (Houston *et al*., 2012). Pallilysin is both an adhesin and a metalloprotease, capable of specifically binding and degrading host ECM components, such as laminin, the most abundant structural component of basement membranes, and fibrinogen, a major component of the host coagulation cascade, as well as fibrin clots, which serve to contain bacteria (Cameron *et al*., 2005; Houston *et al*., 2011, 2012). In addition, heterologous expression of pallilysin in *T. phagedenis* conferred on this non-adherent and non-invasive bacterium the ability to bind to laminin and degrade fibrin clots (Cameron et al., 2008a; Houston et al., 2012).

Thus, I hypothesize that pallilysin is integral to the process of *T. pallidum* dissemination and interference with its functioning will prevent spread throughout the host and establishment of chronic disseminated infection. The objective of my thesis

research was to investigate whether heterologous expression of pallilysin in *T. phagedenis* confers upon this bacterium the ability to cross an artificial endothelial barrier.

1.6 Experimental Approach

A bioinformatics approach was undertaken in an attempt to investigate the evolutionary history of pallilysin in order to gain further insight into its role in *T. pallidum* dissemination. The pallilysin sequence was analyzed in the context of its close and more distant homologues. Several algorithms were used to analyze the degree of conservation of the full length sequence as well as that of specific motifs important for the functioning of the protein. The amino acid changes in the HEXXH motif, as well as in the specific amino acid sequences implicated with laminin binding, were then mapped on a phylogenetic tree in an attempt to track the evolution of these motifs into their current roles in *T. pallidum*.

To investigate whether pallilysin facilitates crossing of the host endothelial barriers by *T. pallidum* in order to establish a disseminated infection, a transwell dissemination assay was developed. First, artificial endothelial barriers were established on permeable filters. Briefly, primary Human Umbillical Vein Endothelial Cells (HUVEC) were grown to confluence on the permeable filters pre-coated with a thin layer of Matrigel matrix, which is a reconstituted basement membrane made up of the same constituents found in *in vivo* basement membranes. The artificial endothelial barriers were suspended in wells, such that growth medium was found on either side of the monolayer. This arrangement made possible the assessment of the intercellular junction integrity of the HUVEC monolayer by measuring the transendothelial electrical

resistance. The presence of upper and lower chambers also allowed quantitation of bacterial traversal across the artificial endothelial barrier down a chemotactic gradient. Percent traversal was assessed to investigate whether heterologous expression of pallilysin confers upon *T. phagedenis* the capacity to disseminate. Dark-field microscopy counts as well as real-time quantitative PCR (QPCR) methodologies were used to establish the percent traversal of *Treponema phagedenis* heterologously expressing pallilysin in comparison to wildtype *T. phagedenis* and *T. phagedenis* transformed with the empty shuttle vector for *pallilysin*, pKMR.

Chapter 2: Materials and Methods

2.1.1 Bacterial Strain and Growth Conditions

Treponema phagedenis biotype Kazan was previously transformed with either the shuttle plasmid pKMR4PEMCS (pKMR) or pallilysin-expressing construct *pallilysin*/pKMR4PEMCS (*pallilysin*/pKMR) (Cameron et al., 2008a). Bacteria were grown at 37°C in an atmosphere of 98% N_2 and 2% H_2 in a custom Coy Laboratory Products anaerobic chamber (Mandel Scientific Company Inc., Guelph, ON) in tryptone-yeast extract-gelatin-volatile fatty acids-serum (TYGVS) medium (Ohta *et al*., 1986).

The TYGVS medium was composed of two solutions. Solution A comprised: tryptone (Difco Laboratories, Detroit, MI), 1.0% w/v; veal heart infusion broth, 0.5% w/v ; yeast extract, 1.0% w/v; gelatin, 1.0% w/v; $(NH_4)_2SO_4$, 0.05% w/v; $MgSO_4$- $_7H_2O$, 0.01% w/v; K_2HPO_4, 0.113% w/v; KH_2PO_4, 0.09% w/v; and NaCl, 0.1% w/v. Solution B was composed of glucose, 0.1% w/v; cysteine hydrochloride, 0.1% w/v; thiamine pyrophosphate, 0.00125% w/v; sodium pyruvate, 0.025% w/v; acetic acid, 0.027% v/v; propionic acid, 0.010% v/v; n-butyric acid, 0.0064% v/v; n-valeric acid, 0.0016% v/v; isobutyric acid, 0.0016% v/v; isovaleric acid, 0.0016% v/v; DL-methylbutyric acid, 0.0016% v/v. The pH of solution B was adjusted to pH 7.2 with 7N KOH; the medium was filter-sterilized, and added aseptically to the autoclaved solution A. Heat-inactivated (30 minutes at 56°C) normal rabbit serum (HI-NRS) (Life Technologies, Frederick, MD) was added to a concentration of 10% to form the complete TYGVS medium.

For culture maintenance, wildtype *T. phagedenis*, as well as pKMR- and *pallilysin*/pKMR/*Treponema phagedenis*, were passaged *in vitro* by inoculating 0.4 ml of

two day-old culture into 10 ml TYGVS medium in the presence of 10 µg/ml rifampicin and 40 µg/ml erythromycin, except for wildtype *T. phagedenis* which was grown in the presence of rifampicin only.

2.1.2 Umbilical Vein Endothelial Cell Line and Monolayer Growth Conditions

Primary Human Umbilical Vein Endothelial Cells (HUVEC) from pooled donors (Lonza, Missisauga, ON) were maintained in EGM™-2 BulletKit™ (Lonza) comprised of basal medium supplemented with hydrocortisone, hEGF, VEGF, hFGF-B, R3-IGF-1, ascorbic acid, heparin, and 2% heat-inactivated fetal bovine serum (HI-FBS). The cells were maintained at 37°C in a 5% CO_2 humidified incubator (Thermo Fisher Scientific, Waltham, MA). HUVEC manipulations were performed in a Biosafety Cabinet Labgard Class II, Type A2 (BSCII) graciously provided by Dr. Terry Pearson. After thawing, cells were grown on tissue culture-treated T75 vented flasks (Cellstar Greiner Bio-One, Monroe, NC) with no additional coating and plated for dissemination assay at passage 3. Cells were thawed at passage 1 and the EGM™-2 BulletKit™ media was changed (25 ml) the following day. HUVECs were grown for two more days and passaged according to Lonza protocol, followed by another passage in two more days immediately before seeding on plates. Media was removed. Cells were washed with 15 ml HEPES-Buffered Saline Solution (HEPES-BSS, Lonza) and trypsinized with 6 ml Typsin/EDTA (Lonza) for 2-3 minutes at room temperature. The solution was neutralized with 1 2ml Trypsin Neutralizing Solution (TNS, Lonza). The cell suspension was transferred to a 50ml falcon tube. After washing the flask with an additional 3 ml of HEPES-BSS, HUVECs were centrifuged at 220 x *g* for 5 minutes at 25° C. The pellet was resuspended in 3 ml of EGM™-2 BulletKit™ media without Gentamicin and Amphotericin that had been pre-

equilibrated in the 5% CO_2 37°C incubator for 3 hours. Finally, 0.5 ml of the HUVEC cell suspension were seeded onto a new T75 flask containing 25 ml medium pre-equilibrated as described above.

2.2 Artificial Endothelial Barrier

To create an endothelial barrier for assessing the traversal rates of the treponemes, HUVECs were seeded on 8.0-µm-pore-size polyethylene terephthalate (PET) filters pre-coated with BD Biocoat growth factor-reduced (GFR) Matrigel in a 24-well Transwell plate (BD Biocoat Invasion Chamber, Becton Dickinson (BD), Franklin Lakes, NJ). BD Matrigel™ matrix is a reconstituted basement membrane preparation extracted from the Engelbreth-Holm-Swarm (EHS) mouse sarcoma, and composed of laminin, collagen IV, entactin, heparin sulfate proteoglycan (perlecan), TGF-β, epidermal growth factor, insulin-like growth factor, fibroblast growth factor, tissue plasminogen activator, and other growth factors which occur naturally in the EHS tumor (BD Biosciences, 2011). The thin layer of pre-coated GFR Matrigel Basement Membrane Matrix served as an *in vitro* reconstituted basement membrane underlying the HUVEC monolayer. The Matrigel and the HUVEC monolayer occluded the filter pores preventing non-invasive bacteria from crossing the artificial endothelial barrier from the insert compartment into the well.

At passage 3, HUVECs were resuspended in 4 ml of EGM™-2 BulletKit™ media without Gentamicin and Amphotericin that has been pre-equilibrated in the 5% CO_2 37°C incubator for 3 hours and were then seeded onto the permeable inserts with the already rehydrated Matrigel. Briefly, two hours prior to seeding, the transwell plate was allowed to thaw at room temperature for about 20 minutes and then the Matrigel coating was rehydrated with growth medium. In the BSCII, 0.5 ml of pre-equilibrated EGM™-2

BulletKit™ media without Gentamicin and Amphotericin and with HI-FBS increased to 12% were added to both inserts and wells. The plate was then transferred to the 5% CO_2 37°C incubator for 2 hours. Afterwards, the medium was removed from the inserts and wells and 1 ml of the pre-equilibrated EGM™-2 BulletKit™ media (as above) was added to all wells, while 0.4 ml of the same medium was added to all inserts. Finally, 0.35ml of media or HUVEC cell suspension ($1.6x10^5$ cells) was added to Matrigel-only or HUVEC inserts, respectively. Thus, the total volume in the inserts was 0.75 ml. After seeding the HUVECs, the plate was rocked back and forth and left to right several times to disperse the cells evenly on the insert membrane and left on the bench at room temperature for 30 minutes for the cells to settle. Then the plate was incubated overnight at 5% CO_2 37°C until a confluent monolayer with tight junctions had formed.

The HUVEC seeding density was determined via a haemocytometer under an Olympus Inverted Light microscope. Two separate 1/5 dilutions of the cell suspension in 0.2% trypan blue solution (Sigma) were prepared and 10 μl of each were loaded on the haemocytometer for counts. Cell concentration was calculated using the average number of viable cells multiplied by the dilution factor and further multiplied by the haemocytometer conversion factor, 10,000.

2.3 Measurement of HUVEC Barrier Function by TEER

The integrity of the monolayer was assessed by measuring Trans Endothelial Electrical Resistance (TEER) using the EVOM2 Voltohmmeter (World Precision Instruments, Sarasota, FL) equipped with Endohm 6 electrodes chamber (WPI). TEER measures the flow of ions across the monolayer and is thus an excellent measure for the integrity and confluence of a monolayer. The transwell plate was taken out of the 5%

CO_2 37°C incubator and placed in the BSCII for about 30-40 minutes for the temperature of the medium to equilibrate to room temperature for consistent measurements, as resistance is affected by temperature. In the mean time, the electrodes were soaked in 70% ethanol for 10 minutes, then rinsed with tissue culture grade distilled water several times, and conditioned with the medium to be used for measurement for 5 to 10 minutes. Medium was then exchanged with 1.5 ml of fresh medium and inserts were sequentially placed in the electrode chamber. Resistance readings from inserts containing only Matrigel, but no HUVECs, served as a baseline and were subtracted from the resistance readings of inserts containing HUVEC monolayers. The resulting resistance readings (with units of Ohms) were then multiplied by the surface area of the inserts (0.3 cm^2) to generate a value that is independent of the growth area of the permeable insert (Haake and Lovett, 1994a). Monolayers with tight junctions typically formed 18 hours after seeding with TEER readings between 15 and 18 ohmsXcm^2. A resistance of 10-12 ohmsXcm^2 is indicative of a confluent monolayer (Grab *et al*., 2005; Haake and Lovett, 1994a; Lovett *et al*., 1988).

2.4 Dissemination across the Artificial Endothelial Barrier

2.4.1 Assay Medium

After overnight incubation of the HUVEC on the transwell inserts, the integrity of the monolayers was assessed by measuring the TEER. The EGM™-2 BulletKit™ media in the inserts and wells were then exchanged with Dissemination Assay Mixed (DAM) medium. The DAM medium was comprised of 25 mM HEPES (Sigma) from a 10X stock, either 10% HI-NRS (inserts) or 20% HI-NRS (wells) with the remaining volume split in 5 parts: 1 part TYGVS medium not including the HI-NRS, and 4 parts EGM™-2

BulletKit™ with 12% HI-FBS but without Gentamicin and Amphotericin. Due to the different concentration of HI-NRS in the inserts and wells a chemotactic gradient across the artificial endothelial barrier was established for the treponemes.

2.4.2 Treponemal manipulations

For dissemination assay experiments 1 ml of one day-old *T. phagedenis* culture was passaged and grown overnight at 37° C in 10 ml TYGVS medium supplemented with 10% heat-inactivated normal rabbit serum (Sigma) in the presence of 10 µg/ml rifampicin and 40 µg/ml erythromycin under anaerobic conditions. Immediately prior to the dissemination assay, the bacteria were harvested by centrifugation of 6ml of culture at 1000Xg for 5 minutes at 25°C. Pellets were resuspended in 3ml room temperature aerobic DAM media with 10% HI-NRS. The motility of the treponemes was examined under the Dark-field microscope (Nikon H550S, Mississauga, ON) and the bacteria were counted in a Petroff Hauser Sperm Counting Chamber. The concentration was adjusted to $6.7x10^7$treponemes/ml and 0.5 ml of either wildtype *T.phagedenis,* pKMR/*T.phagedenis* or *pallilysin*/pKMR/*T.phagedenis* were added where applicable to 0.25 ml of DAM medium with 10% HI-NRS pre-aliquoted in the plate inserts as well as in the wells without inserts.

2.4.3 Experimental Design

Each experiment was designed according to the dissemination assay plate layout as presented in Table 1.

Table 1. Dissemination Assay Plate Layout

Number of wells	Permeable Insert with GFR Matrigel	HUVECs	*T. phagedenis*
2	Yes	no	no
2	Yes	yes	no
2	Yes	yes	wildtype
2	No	no	wildtype
3	Yes	yes	pKMR/*T.phagedenis*
3	No	no	pKMR/*T.phagedenis*
3	Yes	yes	*pallilysin/*pKMR/*T.phagedenis*
3	No	no	*pallilysin/*pKMR/*T.phagedenis*

The two permeable inserts pre-coated with GFR Matrigel without HUVECs or treponemes served as a baseline for the rest of the TEER readings. The two permeable inserts with HUVECs only seeded on the pre-coated with GFR Matrigel filters were used as controls for the HUVEC monolayer integrity under the growth conditions for the duration of the experiment. The permeable inserts that had both HUVECs and either strain of *T. phagedenis* were used to quantify the number of treponemes that traversed the artificial endothelial barrier. Finally, the wells with treponemes but without permeable inserts or HUVECs were used to quantify the total number of treponemes for each strain at the end of the assay. In addition, they served as controls for strain to strain variation in division rates.

2.4.4 Incubation of treponemes with the HUVECs

The treponemes were incubated with the HUVEC monolayer for 48 hours at 25°C in the anaerobic chamber containing 98% N_2 and 2% H_2. At 30 hours the plate was taken out for TEER measurements and returned to the anaerobic chamber after one hour. At the end of the incubation period, the TEER of all monolayers was measured and medium from all the wells was collected. A fraction, 45 µL, was used for dark-field microscopy counts and the rest, 955 µL, was used for bacterial genomic DNA extractions.

2.5 Quantification of Percent Traversal

The percent traversal of the treponemes was calculated as follows:

% Traversal = (Number of treponemes that traversed the artificial endothelial barrier/ Total number of treponemes at the end of the experiment) X 100, where the total number of treponemes represented the number of bacteria from the wells without permeable inserts. The number of bacteria in each well was determined by either dark-field microscopy counting or real-time quantitative PCR (QPCR).

2.5.1 Dark-field Microscopy

Dark-field microscopy counts were performed by counting 5 fields of view of triplicate samples, 12 µL from each well. Appropriate dilutions were prepared from all wells without inserts accounting for the total number of bacteria (1/20), and the wells with inserts containing either pKMR/*T. phagedenis* or *pallilysin*/pKMR/*T.phagedenis* (1/10). No dilutions were made from the wells with inserts containing wildtype *T. phagedenis.*

2.5.2 Genomic DNA extraction

All centrifugation steps were performed at room temperature, at a speed of 12 000 x *g*. Bacteria were harvested from the media obtained from the transwells and the wells without inserts by centrifugation for 10 minutes. The Qiagen DNeasy® Tissue Kit (Qiagen, Toronto, ON) was used for purification of genomic DNA from *T. phagedenis* with a modified extraction protocol. Pellets were resuspended in 180 µl home-made resuspension/lysis buffer (0.5% SDS, 0.1M EDTA, 10mM TRIS, pH 8) and 20 µl proteinase K was added to each tube. Samples were mix thoroughly by vortexing, and incubated at 56°C. Following overnight incubation, samples were vortexed, and allowed to cool down to room temperature for 30 min. Four microliters of RNase A (100 mg/ml) was added to each sample, mixed by vortexing, and incubated for 2 min at room temperature. Samples were vortexed for 15 s and 200 µl Buffer AL was added to each sample, followed by mixing thoroughly by vortexing. Samples were then incubated for 10 minutes at 70°C and mixed by vortexing. Then 200µl ethanol (95%) was added to each sample and mixed again thoroughly by vortexing. The mixtures were then pipetted (including any precipitate) into the DNeasy Mini spin columns placed in 2 ml collection tubes. Columns were centrifuged for 1 min. The DNeasy Mini spin columns were first washed with 500 µl Buffer AW1 and then with 500µl Buffer AW2. An additional 1 minute centrifugation step was performed in new collection tubes to remove any residual ethanol. The DNeasy Mini spin columns were then placed in clean 1.5ml microcentrifuge tubes and dried for 5 minutes at room temperature. Then 100µl PCR Grade MilliQ (Invitrogen) distilled water was added directly onto the DNeasy membrane and incubated at room temperature for 10 min, followed by centrifugation for 2 mins to elute the genomic DNA. A second elution step in a separate tube was performed.

2.5.3 Real- Time Quantitative PCR

The initial copy numbers for *flgE,* a *T. phagedenis* biotype Kazan 5 flagellar hook polypeptide gene (Genbank Accession U04619), were determined using the dsDNA binding dye SYBR Green I. Quantification was performed using a standard curve prepared from genomic DNA extracted from serial dilutions of treponemal cultures. The initial *flgE* copy number was directly proportional, in a 1:1 ratio, to the number of bacteria, from which the amplified DNA was extracted. Thus, by quantifying the initial *flgE* copy number for each sample, the number of bacteria found in the corresponding sample was calculated. The results were then used to assess the percent traversal of each strain.

Standard Curve

T. phagedenis culture was grown to mid-exponential stage and 10-fold serial dilutions were prepared in triplicate ranging from $1x10^8$ to $1x10^2$ treponemes/ml each in volumes of 1ml. The samples were spun down at 12, 000Xg for 10 min at room temperature and genomic DNA was extracted. A region in the *flgE* gene was amplified and a standard curve was created by plotting initial *flgE* copy number in the genomic DNA of each dilution against the resulting threshold cycle (Ct) value. Samples were run in quadruplicate.

PCR Conditions

The size of the resulting amplicon was 245 bp. The sequence of the primers for the target gene were FlgE 9q F 5' GAGGGGCAGCAAATAATCAA 3' and FlgE 9q Rii 5' GGGACCCTTGAGAAAACCAT 3'.

The PCR reactions were performed on an Eppendorf Mastercycler Realplex 4 using white 96-well qPCR plates (Eppendorf twin-tec). Reaction mixtures (20 µL) for the quantification of *flgE* copy number in the dissemination assay samples contained template DNA, 1X in-house-made SYBR Green I buffer (0.05 M Tris-HCl, 0.05 M Tris, 0.2 M KCl, 8% Glycerol, 0.1% Tween-20, 2.5x SYBR Green I, 0.05 M Ammonium Sulphate), 4 mM MgCl2, 200µM dNTPs (Thermo Scientific), 500 nM of each primer, and 0.8 Units Maxima Hot Start Taq DNA Polymerase (Thermo Scientific). Each PCR run included: 1) test samples (in quadruplicate), 2) one to four positive controls from the standard curve (dilutions $1x10^5$, $1x10^6$, $1x10^7$, and $1x10^8$ treponemes/ml, in triplicate), 3) a no-template control (in triplicate), 4) a no-primer control (in triplicate), and 5) a no-amplification control (in triplicate), where no polymerase was added. The cycling conditions comprised 10 minutes at 95°C, 40 cycles at 95°C for 15 seconds, 65°C for 20 seconds, and 72°C for 40 seconds.

Melting Curve Analysis of SYBR Green I Assays

For each PCR run with SYBR Green I detection, a melting curve analysis was performed to verify the specificity of amplification for each reaction, i.e. absence of primer dimers and other nonspecific products. Briefly, reactions were heated to 95°C for 15 seconds, cooled down and maintained at 60°C for 15 seconds, and melting curve analysis was performed over 20 minutes of continuous heating to 95°C.

2.6 Confirmation of Strain Identity

Plasmid DNA Extraction

Following the dissemination assay, plasmid DNA from *pallilysin*/pKMR/*T. phagedenis* and pKMR/ *T. phagedenis* was extracted and sequenced to verify the identity of the two strains. The cultures were grown for 3 days, then 1ml was inoculated into 10 ml TYGVS with 10% HI-NRS. Following a 24-hour incubation period, the bacteria were harvested by sequential centrifugation of a total of 8ml of culture per strain. Briefly, 4 ml were aliquoted first in 1ml fractions into four 1.5ml microcentrifuge tubes, spun down at 8000 x *g* for 2 mins at room temperature, and the supernatant was removed. The other 4 ml were then added to the pellets in the same manner. The Thermo Scientific GeneJET Plasmid Miniprep Kit was then used to extract plasmid DNA with a modified extraction protocol. Each tube was treated as one "sample" for alkaline lysis, i.e. 2 ml culture material was resuspended in 250 µl resuspension solution, 250 µL of the Lysis Solution were added and mixed thoroughly by inverting the tubes 6 times, and then 350 µL of the Neutralization Solution were added and mixed immediately and thoroughly by inverting the tube 6 times. All centrifugation steps were performed at room temperature at 12 000 x *g*. Samples were centrifuged for 10 minutes to remove the precipitate. The supernatants were sequentially combined onto an individual GeneJet™ spin column, and columns were washed twice with 500 µl Wash Solution. The columns were re-centrifuged in new collection tubes for 1 minute to remove any remaining wash buffer. Columns were then transferred to 1.5 ml microcentrifuge tubes and 30 µl room temperature PCR grade MilliQ (Invitrogen) distilled water was added directly to the column. Following incubation for 2 minutes at room temperature, the tubes were centrifuged for 2 minutes.

The elution step was repeated in the same microcentrifuge tube using the eluted plasmid solution to increase the concentration of plasmid DNA.

Plasmid DNA Sequencing

The primers, pKMR 3' seq and pKMR 5' seq, amplified a region surrounding the MCS where *tp0751* is inserted. The sequences for the sequencing primers were: pKMR-5'-seq: 5' GTT ATC GTA TTA TTT AAC GGG A 3'; pKMR-3'-seq: 5' GTG GCC ACT AGT ACT TCT 3'. Each tube containing 8µl pDNA and 4µL either forward or reverse primer was sent for sequencing at Eurofins MWG Operon (Huntsville, AL).

2.7 Statistical Analysis

Due to time restrictions with a long period of optimizations and issues with contamination, statistical analysis was performed on the results from a single plate rather than the originally intended results from three replica plates. Although the statistical requirement for independent samples was not met, since the analysis was performed on a single plate, each replica transwell on the plate was treated as an independent sample as a proof of principle. Each *T. phagedenis* strain was considered a treatment with the number of replica transwells of each treatment as follows: wildtype *T. phagedenis* (n=2), pKMR/*T. phagedenis* (n=3), and *pallilysin*/pKMR/*T. phagedenis* (n=3). Analysis was performed on Graph Pad Prism 6 (La Jolla, CA) using one-way ANOVA and $\alpha=0.05$ to determine if the differences between the treatments' means were greater than expected to have occurred by chance. Tukey's multiplicity test was used to compare the means of each treatment to the means of every other treatment. Power analysis was subsequently carried out with StatMate software to assess the probability of detecting significance with

the current experimental design. Sample size calculations were then performed to estimate the minimal sample size needed to obtain significance (α=0.05) with power of 80%.

2.8 *In vitro* pallilysin Activation Assay

The *E. coli*-expressed recombinant pallilysin proteins (wild-type HAXXH and mutant HEXXH, E199Q), (30 µg) were pre-incubated at 37°C for 21 h and mixed with 60 µg of plasminogen-free human fibrinogen (Calbiochem, Darmstadt, DE) in a HEPES-based protease activation buffer (20 mM HEPES, pH 7.0, 25 mM $CaCl_2$). The reaction mixtures, as well as fibrinogen alone, were incubated at either 37°C or 25°C for 48 hours, and samples were removed every 24 hours. Samples from each time point (100 µl) were mixed with 100 µl of SDS-sample buffer with DTT and heated at 95°C for 10 min, and the three fibrinogen chains, α, β, and γ, were separated and analyzed for degradation by electrophoresis in 5% stacking and 15% separating SDS polyacrylamide gels, at a constant voltage of 200 V for 1 h. Gels were stained in Coomassie brilliant blue R-250 solution and destained in 5% v/v acetic acid, 20% v/v methanol, and 75% v/v H_2O.

2.9 Bioinformatics analyses

Protein/nucleotide sequence information

Protein and nucleotide sequence information for pallilysin was obtained using the Microbial Genome Database (MBGD) (Uchiyama, 2003) and the National Center for Biotechnology Information (NCBI) GenBank database (Benson *et al.*, 2003). The strain used for all *T. pallidum* sequence information was Nichols (Fraser *et al*., 1998). The strain used for all *T. phagedenis* sequence information was Kazan.

Identifying *Treponema pallidum* pallilysin homologues

The NCBI database was searched for homologues of pallilysin using protein Basic Local Alignment Search Tool (BLASTp) and Position-Specific Iterated BLAST (PSI-BLAST) analyses (Altschul *et al.*, 1990). A multiple sequence alignment based on the BLASTp analysis was created with ClustalW (Larkin *et al*., 2007). A profile hidden Markov model (HMM) (Eddy, 1998, 2010) was created from the multiple sequence alignment using HMMER (Finn *et al*., 2011) and used to search the UniProtKB database for additional homologs of pallilysin.

Analysis of protein sequences

The pallilysin homologues were further analyzed using MEGA 4 (Tamura *et al*., 2007). Amino acid composition and frequencies were assessed in the pallilysin sequence and in comparison to the rest of the homologues. The phylogenetic relationship of the homologues was estimated on Neighbour-joining distance trees (Gascuel and Steel, 2006a; Morrison, 1996). The evolutionary distances were computed using the Dayhoff substitution matrix (Dayhoff *et al*., 1978). The reliability of the phylogenetic estimation was assessed using the bootstrap method (Felsenstein, 1985; Soltis and Soltis, 2003).

Sequence logos of the multiple sequence alignments were created using Weblogo to search for patterns in sequence conservation. Based on the sequence information from the multiple sequence alignments and the WebLogos, a survey of conserved and variable amino acids was carried out to gain further insight into the evolutionary history and function of the pallilysin sequence. In addition,specific sequence changes in the laminin-

binding and HEXXH motifs among the pallilysin homologues were mapped on the phylogenetic tree.

Chapter 3: Bioinformatics Analysis on the *Treponema pallidum* Protease, Pallilysin

3.1 Contributions to the data

My contributions to the data presented in this chapter include: BLASTp, PSI-BLAST, and HMMER search analysis of the *T. pallidum* protease, pallilysin; multiple sequence alignments of the BLASTp and HMMER hits, creation of WebLogo sequence logos and neighbour-joining phylogenetic trees with selected sequence changes mapped onto it. This work was done in collaboration with Dr. John Taylor at the University of Victoria, Canada.

3.2 Introduction

Bioinformatics involves the application of computational techniques to understand and organise the information associated with biological macromolecules (Luscombe *et al*., 2001). The use of bioinformatics in molecular biology research includes, but is not limited to, sequence alignment, gene identification, genome assembly, drug design and discovery, protein structure prediction, prediction of gene expression and the modeling of evolution (Bansal, 2005). One of the driving forces behind bioinformatics is the search for similarities between different biomolecules (Luscombe *et al*., 2001).

Computer analyses of the *T. pallidum* genome were previously performed to identify ORFs predicted to encode outer membrane proteins, as such proteins may be surface exposed and thus are likely to be involved in host cell attachment (Cameron,

2003). Ten ORFs were expressed based on strong PSORT predictions for potential outer membrane location, predicted structural similarity to known beta-barrel-containing proteins (3DPSSM analysis), or presence of motifs characteristic of outer membrane proteins (PROSITE analysis). Attachment studies showed that recombinantly expressed pallilysin (Tp0751), one of the ten expressed ORF, was a laminin-binding adhesin (Cameron, 2003). Additional studies using synthetic peptides localized the minimum laminin-binding region to 10 amino acids: P^{98}, V^{99}, Q^{100}, T^{101}, W^{127}, I^{128}, T^{182}, A^{183}, I^{184}, and S^{185}(Cameron *et al*., 2005). Further bioinformatics analyses of the primary protein sequence of pallilysin revealed the presence of an HEXXH motif (HEVIH) at the C-terminus the protein, amino acids H^{198} to H^{202} (Houston *et al*., 2011). Such a motif is found in approximately 50% of metalloproteases and was shown to be an essential part of the active site of pallilysin responsible for the activation and proteolytic function of the protein (Houston *et al*., 2011, 2012).

To further characterize pallilysin, in the current study bioinformatics analyses of the amino acid sequence were performed in the context of multiple sequence alignments of homologous protein sequences from other bacteria. Generation of sequence logos and phylogenetic trees based on the multiple sequence alignments allowed assessment of the evolutionary conservation of the pallilysin sequence. Finally, two motifs in the pallilysin sequence important for dissemination of *T. pallidum,* the HEXXH and the core laminin-binding motifs, were mapped onto the phylogenetic tree in the context of the pallilysin homologues. The analysis of the conservation of these and other amino acids from the pallilysin sequence served to infer the evolutionary history and role of pallilysin in a widely disseminating pathogen.

3.3 *Treponema pallidum* pallilysin (Tp0751)

T. pallidum gene *tp0751* (NC_000919.1) is located at position 815821..816534 on the chromosome (Fraser, 1998). The gene product, pallilysin (Tp0751), is a 237 amino acid protein.

Basic Local Alignment Search Tool (BLAST) search of the pallilysin protein sequence

The essential assumption in any phylogenetic analysis is that all the sequences are homologous, that is, descended from a common ancestor (Hall, 2013). Alignment programs will align sequences, homologous or not, and tree-building programs will make a phylogenetic tree from that alignment. However, the tree would be misleading unless the sequences were homologous. A reliable way to identify sequences that are homologous to the sequence of interest is to do a Basic Local Alignment Search Tool (BLAST) search (Altschul *et al*., 1997) with the sequence of interest as a query. The program identifies regions of local similarity between sequences and calculates the statistical significance of matches represented by the E-value. The lower the E-value is, the lower the probability that the match occurred due to chance.

BLASTp was used to compare the amino acid sequence of *T. pallidum* pallilysin to all other sequences found in the NCBI database to infer functional and evolutionary relationships between the sequences. The outcome was a limited number of BLASTp hits listed in Table 2 in decreasing order of the absolute E-value. The corresponding percent identity to pallilysin was also included. The table also lists the bacterial order and species from which the homologous sequences originated.

Table 2. BLASTp results for *Treponema pallidum* protein pallilysin.

Protein	Bacterium	Sequence ID	% Identity	E-value
Pallilysin (Tp0751)	*Treponema pallidum* subsp. *pallidum*	NP_219188.1	237/237(100%)	2e-171
laminin-binding protein	*Treponema paraluiscuniculi*	YP_004673471.1	234/237(99%)	1e-168
HP	*Treponema phagedenis*	WP_002695045.1	50/120 (42%)	4e-22
HP	*Treponema vincentii*	WP_006187550.1	44/116 (38%)	6e-18
HP	*Treponema medium*	WP_016523557.1	45/114 (39%)	9e-18
Tde0840	*Treponema denticola*	NP_971450.1	39/125 (31%)	2e-14
HP Tpe0045	*Treponema pedis*	YP_008433684.1	35/120 (29%)	3e-12
HP Trebr0632	*Treponema brennaborense*	YP_004439205.1	34/120 (28%)	2e-09
HP Spica1139	*Treponema caldaria*	YP_004697793.1	36/113 (32%)	3e-09
HP Treaz3334	*Treponema azotonutricium*	YP_004528131.1	39/123 (32%)	4e-09
HP Tresu1502	*Treponema succinifaciens*	YP_004365699.1	32/105 (30%)	9e-09
HP	*Treponema lecithinolyticum*	WP_021686449.1	36/130 (28%)	3e-08
HP	*Treponema maltophilum*	WP_016526295.1	36/132 (27%)	1e-07
HP	*Treponema socranskii* subsp. *socranskii*	WP_021329513.1	38/145 (26%)	1e-07
OmpA/MotB domain protein	*Treponema saccharophilum*	WP_002702225.1	31/120 (26%)	2e-07
HP Trepr1835	*Treponema primitia*	YP_004531174.1	35/115 (30%)	2e-07
HP	*Treponema socranskii* subsp. *paredis*	WP_016521899.1	32/126 (25%)	1e-06
HP	*Treponema* sp. *JC4*	WP_009107102.1	30/105 (29%)	0.001
HP	*Spirochaeta thermophila*	YP_006045013.1	29/122 (24%)	0.031
HP	*Treponema bryantii*	WP_022932903.1	27/117 (23%)	1.2
HP	*Spirochaeta africana*	YP_005475377.1	33/110 (30%)	8.2

HP: Hypothetical proteins

The matches were almost exclusively derived from the genus *Treponema* with the exception of two, which came from the genus *Spirochaeta*. It is important to note that the E-values of the last four members in the table suggested possible alignment by chance rather than homology; however, they were kept in the following analyses in attempt to include more distant homologues of pallilysin. Both the *Treponema* and *Spirochaeta* genera belong to the *Spirochaetaceae* family (Whitman, 2010). Based on the BLASTp search, it appeared that the pallilysin amino acid sequence was conserved primarily within the *Treponema* genus and to a smaller extent in the *Spirochaeta* genus.

Highest percent identity (99%) was observed with the sequence for the laminin-binding protein from *T. paraluiscuniculi*, which has reportedly lost its pathogenicity to humans, but causes venereal syphilis in rabbits (Centurion-Lara *et al*., 2013). The sequence with the second highest (42%) percent identity to pallilysin was that of *T. phagedenis.* As mentioned earlier, *T. phagedenis* was originally isolated from a syphilitic chancre and was infectious to rabbits, but is no longer considered pathogenic to the latter. In addition, it is believed to be a commensal of the human flora (Trott *et al*., 2003; Wallace and Harris, 1967). It is puzzling why such high percent identity exists between these two sequences. Likely the answer lies in the once pathogenic nature of *T. phagedenis* (Trott *et al*., 2003; Wallace and Harris, 1967). Further analysis of specific motifs will be presented later in this chapter. Two sequences of lower percent identity (38 and 39%) from two human oral pathogens, *T. vincentii* and *T. medium*, followed closely. Dashper *et al*. (2011) argues that while treponemes from the oral cavity are most often associated with periodontal diseases, these treponemes are members of the normal oral microbiota of healthy individuals in very low numbers. It is possible that the case is

similar with *T. phagedenis.* The rest of the BLASTp hits comprised sequences of percent identities equal to or lower than 32% and originated from both free-living and host-associated pathogenic or commensal species. The pathogenic species from the list of sequence matches, with the exception of *T. paraluiscuniculi,* were all associated with localized infections, such as bovine digital dermatitis and gingivitis or periodontal disease in people (Dashper *et al*., 2011; Evans *et al*., 2009). Infection with *T. paraluiscuniculi* or *T. pallidum,* however, results in a systemic disease (Lafond and Lukehart, 2006). *T. pallidum* and *T. paraluiscuniculi* have a higher dissemination capacity than the rest of the pathogenic treponemes and, interestingly, the pallilysin homologues from these two species share the highest extent of sequence identity.

Position-Specific Iterative Basic Local Alignment Search Tool (PSI-BLAST) search of the pallilysin protein sequence

The Position-Specific Iterative Basic Local Alignment Search Tool (PSI-BLAST) was performed to look for more divergent homologous sequences beyond what BLASTp discovered. PSI-BLAST is a multi-iteration search tool used to identify distant homologues in the protein database (Altschul *et al*., 1997; Bhagwat and Aravind, 2007; Eddy, 1998). It starts with a single query sequence and identifies homologous sequences using BLASTp. However, while BLASTp uses position-independent scoring parameters, PSI-BLAST derives a position-specific scoring matrix or a profile from the multiple sequence alignment of the homologous sequences. The matrix identifies the conservation pattern in the alignment and transforms it into a matrix of scores for each position in the alignment. The profile is then searched against the database to find more distant

homologues that match the conservation pattern. The profile is updated with each successive iteration. Therefore, PSI-BLAST is more powerful at detecting distant sequence similarities than BLASTp is. Rather than relying on a single sequence and position-independent scoring parameters, PSI-BLAST utilizes a score matrix that contains the conservation information from an alignment of related sequences. Upon two consecutive iterations of PSI-BLAST using pallilysin as the query sequence, additional sequences appeared of less than 20% identity. Those included sequences of hypothetical proteins from *Spirochaeta alkalica, Spirochaeta smaragdinae, Spirochaeta bajacaliforniensis* and *Spirochaeta* sp. *L21-RPul-D2*. On the third iteration, sequences from borrelia species were included with percent identities ranging from 19% down to 12%. Some of these sequences belonged to *B. hermsii,* the causative agent of relapsing fever. Despite the low percent identity, the E-values suggested that these hits were likely homologous in sequence to the pallilysin sequence. It is possible that motifs or whole domains were added or deleted throughout the course of evolution of these sequences in their respective organisms. Such divergence is also likely to be related to the particular functions of these proteins. Nonetheless, based on the PSI-BAST searches it appeared that the conservation of the pallilysin sequence was not limited only to the *Treponema* and *Spirochaeta* genera, but also to a lesser extent to the *Borrelia* genus. Unfortunately, the hits were not annotated and remained as hypothetical proteins in the database. However, the phylogenetic scope of the pallilysin homologues and the extent of sequence conservation of pallilysin as a whole and with respect to particular functionally important domains could provide supporting evidence for the role of the protein in *T. pallidum* dissemination.

HMMER search of the pallilysin protein sequence

To confirm and build upon the results from PSI-BLAST, another probabilistic methods search tool, HMMER, was utilized to search the UniProtKB database for homologues of pallilysin. HMMER makes a profile of the query sequence alignment that assigns a position-specific scoring system for substitutions, insertions, and deletions (Finn *et al*., 2011). Unlike the profiles created with PSI-BLAST, HMMER profiles are probabilistic models called profile hidden Markov models (profile HMMs) (Eddy, 1998, 2010). Profile HMMs are statistical models of the consensus of the multiple sequence alignment. They capture position-specific information about how conserved each column of the alignment is, and the varying degree to which gaps and insertions are allowed. The Position Specific Scoring Matrix of PSI-Blast differs from the profile HMMs used in HMMER in that gaps are placed according to a fixed penalty plus a penalty for extension, rather than learning position specific gap penalties from the sequences already aligned (Altschul *et al*., 1997; Finn *et al*., 2011). Therefore, HMMER has the potential to find homologous sequences that PSI-BLAST had missed. For this reason, a multiple sequence alignment of the BLASTp results from Table 2 was created using ClustalW. The resulting sequence alignment was then used as a query in HMMER to create an HMM profile and search the UniProtKB database for homologues of pallilysin. The results were similar to those of PSI-BLAST (3^{rd} iteration), but also included other borrelia species, namely the etiological agent of Lyme disease, *B. burgdorferi.* The output from HMMER was used to create a second multiple sequence alignment. The alignment was then used to construct a phylogenetic tree and a sequence logo.

Multiple Sequence Alignments

The BLASTp sequences from Table 2 and the HMMER output sequences were imported into BioEdit, a biological sequence alignment editor (Hall, 1999), and multiple sequence alignments were generated through ClustalW (Larkin *et al*., 2007). The program calculates the best matches for the selected sequences, and lines them up so that the identities, similarities and differences can be seen.

Analysis of the individual lengths of the sequences in the multiple sequence alignments was performed in Mega 4 (Tamura *et al*., 2007). The size comparison revealed that both the sequence of pallilysin and that of its homologue from *T. paraluiscuniculi* were equal in length (237 amino acids). It should be noted that the latter has been annotated as a laminin-binding protein due perhaps to the high sequence similarity with pallilysin and will henceforth be referred to as pallilysins's laminin-binding homolog from *T. paraluiscuniculi*. Both homologues were about half the size of the rest of the homologues, with the exception of the 800-amino-acid-long homologue from *T. saccharophilum*. Next, amino acid composition analyses were performed based on the full-length BLASTp and the HMMER alignments. The analyses provided clues about the frequency of incorporation of certain amino acids in the pallilysin sequence in comparison to the amino acid frequencies in the rest of the homologues. In dealing with amino acid frequencies, the length of each homologue is irrelevant, and therefore comparison of amino acid frequencies among homologues of different sizes is possible. Thus, amino acid frequency ratios were calculated by dividing the frequency of each amino acid in the pallilysin sequence by its averaged counterpart in the multiple sequence

alignment. Thus, a ratio greater than 1 for a particular amino acid signified increased frequency of that amino acid in the pallilysin sequence in comparison to the rest of the homologues. A ratio lower than 1 implied decreased amino acid frequency in pallilysin, and a ratio equal or close to 1 suggested no difference in amino acid frequencies. As seen in Table 3, the amino acids alanine, histidine, and proline were found in much higher frequencies in the pallilysin sequence than in the rest of the sequences from the multiple alignments. Conversely, the amino acids glutamate, isoleucine, lysine, asparagine, and tyrosine were found in much lower frequencies in the pallilysin sequence.

Table 3. Analysis of pallilysin amino acid frequencies.
The amino acid frequencies in pallilysin and the rest of the homologues were compared. The HMMER output included additional *Spirochaeta* and *Borrelia* homologues not present in the BLASTp output. All frequencies were presented in percentages, and the ratios were obtained by dividing the pallilysin amino acid frequencies by the average amino acid frequencies from all the homologues in the alignments.

		Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu
HMMER	Pallilysin	12.24	0.42	4.22	3.8	4.22	5.91	6.33	5.06	0.84	7.17
	Average	4.13	0.25	6.22	7.2	5.72	4.52	1.27	9.68	8.1	8.39
	Ratio	3.0	1.7	0.7	0.5	0.7	1.3	5.0	0.5	0.1	0.9
BLASTp	Pallilysin	12.24	0.42	4.22	3.8	4.22	5.91	6.33	5.06	0.84	7.17
	Average	6.16	0.29	5.88	7.32	4.57	5.79	1.15	8.42	5.58	7.76
	Ratio	2.0	1.4	0.7	0.5	0.9	1.0	5.5	0.6	0.2	0.9

		Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr
HMMER	Pallilysin	1.27	2.95	8.02	3.8	7.17	9.28	8.86	5.49	1.69	1.27
	Average	1.39	6.68	2.93	3.91	3.77	8.4	5.28	6.27	0.87	5.06
	Ratio	0.9	0.4	2.7	1.0	1.9	1.1	1.7	0.9	1.9	0.3
BLASTp	Pallilysin	1.27	2.95	8.02	3.8	7.17	9.28	8.86	5.49	1.69	1.27
	Average	1.64	5.08	3.32	4.87	5.14	8.52	7.4	6.09	1.23	3.81
	Ratio	0.8	0.6	2.4	0.8	1.4	1.1	1.2	0.9	1.4	0.3

Based on the calculated amino acid frequency ratios, it is tempting to infer a connection between the higher frequency of histidines in the pallilysin sequence and the HEXXH motif found in the pallilysin active site. Analogous reasoning applies to the laminin-binding protein of *T. paraluiscuniculi* since its histidine frequency ratio is the same as in pallilysin. Perhaps, in the course of evolution, the higher tendency of the pallilysin sequence for having histidines facilitated the formation of this motif and contributed to the role of the protein in the process of *T. pallidum* dissemination. Therefore, the HEXXH could be a motif that gained function, rather than lost one in the divergence of this protein sequence. Further analysis of the amino acid changes of specific motifs in the pallilysin sequence will be considered later in this chapter based on the phylogenetic tree analyses.

An important feature of multiple sequence alignments is the introduction of gaps into the sequences (Hall, 2013). These gaps represent historical insertions or deletions, and they serve to bring homologous sites into alignment in the same column. It is important to keep in mind that an alignment is simply an estimate of the positions of historical insertions and deletions; however, the quality of the alignment can affect the quality of the phylogenetic analysis in a tree. In dealing with the problematic aspect of sequence alignments, it is a common practice to delete parts of the sequences that cannot be aligned reliably for the phylogenetic analysis, because an unreliable hypothesis of homology can lead to poor estimation of phylogenetic reconstruction.

Since the pallilysin query sequence was much shorter than the rest of the homologous sequences, with the exception of the laminin-binding protein from *T. paraluiscuniculi,* and the coverage for the other hits was generally only around 50%, a

great number of gaps were introduced by ClustalW. These gaps were mostly placed in the N-terminal domain of the sequence. Therefore, selective deletion of certain regions from the multiple sequence alignment was performed to avoid potential ambiguous phylogenetic estimation. The remaining region after the deletion of the highly variable flanking regions spanned amino acids L^{122} – Y^{224} in the pallilysin sequence. The selected region contained the HEXXH motif, but only two-thirds of the core laminin-binding amino acid sequence.

Sequence logos of the multiple sequence alignments using Weblogo

Sequence logos were created using WebLogo 2.8.2 based on the multiple sequence alignments from the BLASTp and HMMER searches. Sequence logos are graphical representations of patterns within a multiple sequence alignment of either amino acid or nucleic acid sequences (Crooks *et al*., 2004; Schneider and Stephens, 1990). Each logo consists of stacks of symbols, in which one stack represents each position in the multiple sequence alignment. The overall height of the stack indicates the sequence conservation at that position, while the individual height of each symbols within the stack correlates to the relative frequency of each amino or nucleic acid at that position. A sequence logo displays patterns in sequence conservation and aids in discovering and analyzing such patterns; therefore, it can provide a richer and more precise description of, for example, a functional domain or a motif, than would a consensus sequence.

A WebLogo of the full coverage of the multiple sequence alignment based on the BLASTp search revealed a fair number of highly conserved amino acid residues located

mostly in the mid region of the alignment. In addition, WebLogos of the truncated (L^{122} – Y^{224}) region containing the HEEHX motif were generated from the alignments resulting from both the BLASTp and the HMMER searches (Figure 8). The HEXXH motif occupies amino acids H^{198} – H^{202} in the original pallilysin sequence, and amino acid positions 402 through 406 in the WebLogos. Although the HEXXH motif itself is not fully conserved among the sequences from the multiple sequence alignments, a high level of conservation was apparent in the surrounding regions as can be seen in Figure 7. It seems that the HEXXH motif is unique in *T. pallidum* and *T. paraluiscuniculi*, and might have evolved in these two species in agreement with their systemic pathogenic profile. The flanking sequences, however, contain a number of highly conserved residues, as evident by the height of the WebLogo symbols at their respective positions. Some of the conserved amino acids in the WebLogos of the truncated multiple sequence alignments include: F, G, L, W, D, E, I, N, R, Y. Phenylalanine (F), glycine (G), leucine (L), tryptophane (W), aspartate (D) and arginine (R) are found at similar frequencies in pallilysin and the rest of the HMMER homologues (Table 3). Thus, it is reasonable that these amino acids would have been conserved. However, the glutamate (E), isoleucine (I), and tyrosine (Y) residues are found in much lower frequencies in the pallilysin sequence. Therefore, their conservation at specific positions among the majority of the homologous sequences implies the importance of these amino acids at the specific positions they occupy in the sequences. It is interesting to note that the isoleucine (I^{405} in the WebLogos) found in the HEXXH motif of pallilysin (HEVIH) is highly conserved among not only Treponema, and Spirochaeta, but also among Borrelia species. In a very

few cases a conservative point mutation to valine (V) had occurred, where both amino acids are hydrophobic, non-polar and of similar size.

Taken together, the conservation of the flanking sequences of the HEXXH motif in pallilysin, as well as of the isoleucine in its core, and the higher frequency of histidines in comparison to the other homologues, strongly favor an evolutionary recent gain of function as a metalloprotease, important in dissemination of *T. pallidum* and *T. paraluiscuniculi*. Indeed, the *T. denticola* homologue failed to degrade fibrinogen (Houston *et al*., unpublished observations). Additionally, the degree of conservation among the homologues, indicated by the sequence logos, strongly supports a conserved alternate function in all species.

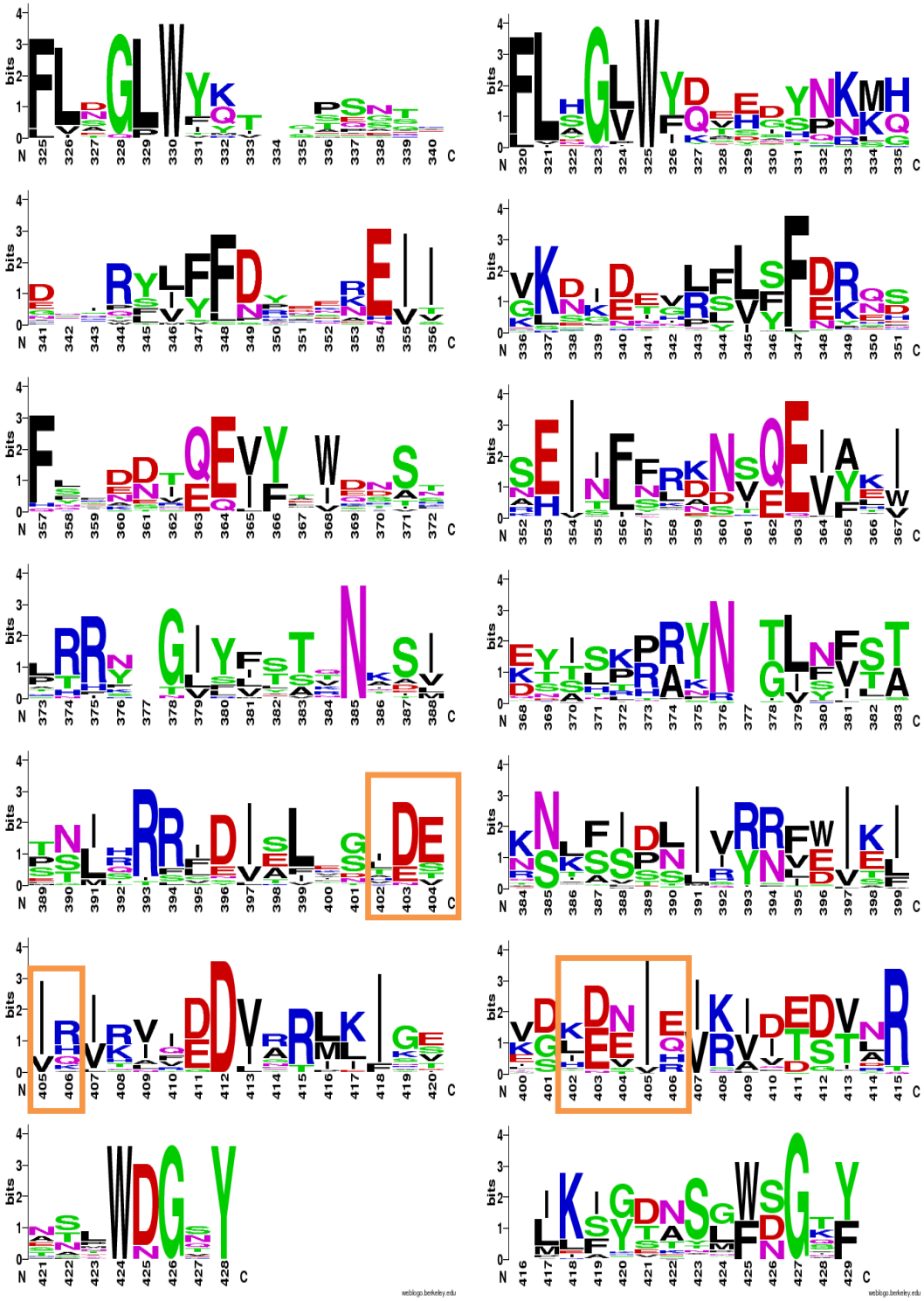


Figure 7. WebLogos showing conservation in the truncated (L^{122} – Y^{224}) pallilysin region containing the HEXXH motif. The WebLogos were based on multiple alignments generated from the BLASTp (left) and HMMER (right) searches. The former contains 21 sequences, while the latter contains 80 sequences and includes homologues from *Borrelia* species. The HEXXH motifs comprise the boxed region of the WebLogos.

Phylogenetic Tree construction

The truncated multiple sequence alignments from the BLASTp and the HMMER searches, used to generate WebLogos, were also used to estimate phylogenetic trees (Figures 9 and 10). The trees were reconstructed using the MEGA 4 (Molecular Evolutionary Genetics Analysis) software, which is an integrated tool for creation and exploration of sequence alignments, the estimation of sequence divergence, the reconstruction and visualization of phylogenetic trees, and the testing of molecular evolutionary hypotheses (Tamura *et al*., 2007). The same program was also previously used to analyse the amino acid composition of the multiple sequence alignments as presented in Table 3.

A phylogenetic tree, itself, is an estimate of the relationships among species or sequences and their hypothetical common ancestors (Hall, 2013). A phylogenetic tree consists of external nodes representing the analyzed sequences, internal nodes representing hypothetical ancestors, and branches that connect the nodes to each other (Hall, 2013; Morrison, 1996). The lengths of the branches represent the amount of change that is estimated to have occurred between a pair of nodes.

The method of phylogenetic analysis chosen for the pallilysin multiple sequence alignments was neighbor-joining. It is an algorithmic method that analyses distance data (Gascuel and Steel, 2006a; Morrison, 1996). It starts by initially grouping the two sequences with the smallest distance, and then progressively adds more distant sequences to that or a new group. Each pair of groups is adjusted on the basis of their average divergence from all other groups. The tree is constructed so that the observed distance

between any two sequences is equal to the sum of the branch lengths connecting them on the tree.

The evolutionary distances were computed using the Dayhoff matrix based method. The Dayhoff matrix is a substitution matrix used to score sequence alignments (Dayhoff *et al*., 1978). Each amino acid substitution is given a score based on the likelihood of the amino acid being replaced with another one during a specified evolutionary interval, rather than these two amino acids being aligned due to chance. The method reflects the chemical meaning of amino acid mutations, in which groups of chemically similar amino acids tend to replace one another to maintain the functionality of specific interactions within the three-dimensional conformation of proteins. Factors that govern such interactions are, for example, size, shape, and local concentration of electric charge, ability to form hydrogen bonds, salt bonds, or disulfide bonds. The Dayhoff matrix is based on substitutions observed throughout the alignment including both highly conserved and highly mutable regions. This method thus allows alignment of more divergent homologous sequences, unlike the identity matrix, in which an amino acid is only maximally similar to itself and substitutions are not allowed. Therefore, the Dayhoff matrix method was chosen to estimate the phylogenetic relationships between pallilysin and its homologous sequences from both the BLASTp and the HMMER searches.

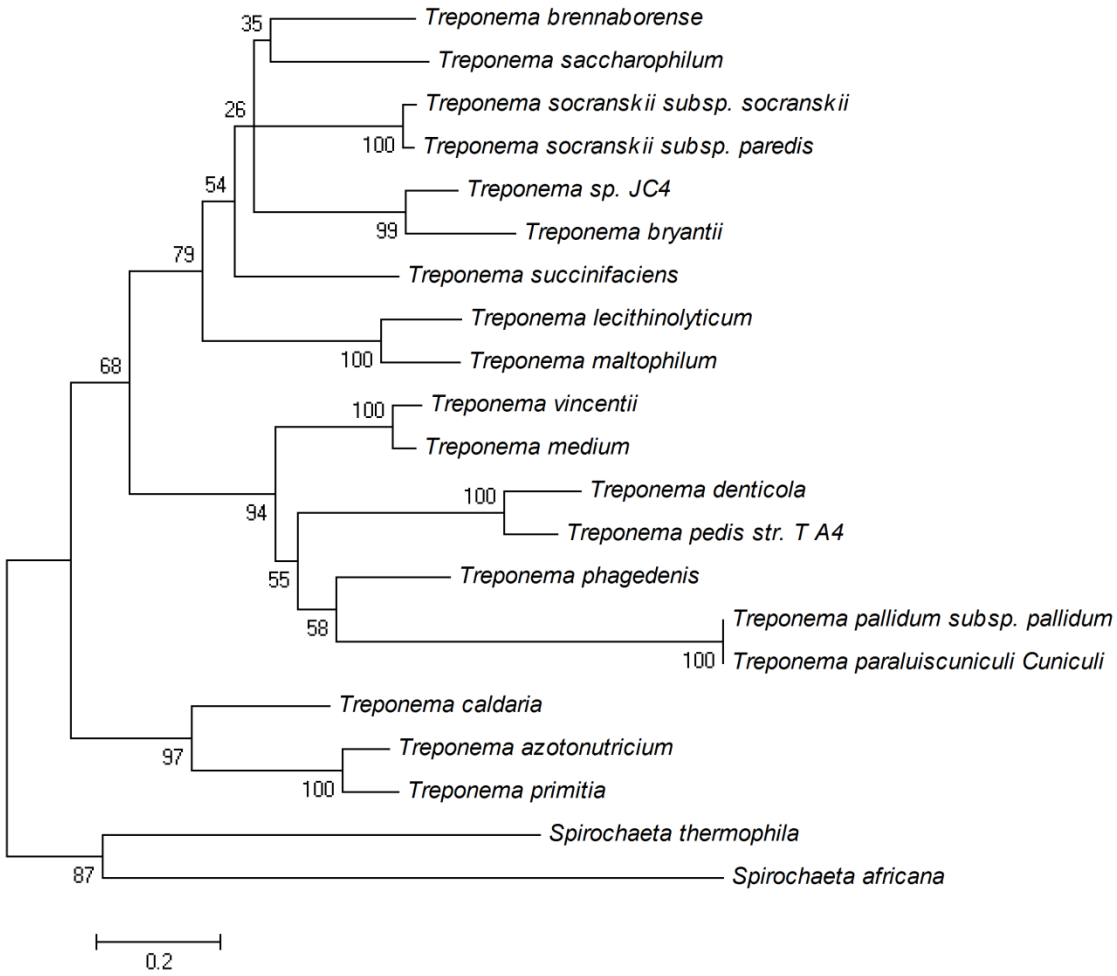


Figure 8. Evolutionary relationships of 21 BLASTp pallilysin homologues
The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 7.14225343 is shown. The percentage of replicate trees in which the associated sequences clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Dayhoff matrix based method and are in the units of the number of amino acid substitutions per site. The analysis involved 21 amino acid sequences. All ambiguous positions were removed for each sequence pair. There were a total of 104 positions in the final dataset. Evolutionary analyses were conducted in MEGA 4 (Dayhoff *et al*., 1978; Gascuel and Steel, 2006b; Soltis and Soltis, 2003; Tamura *et al*., 2007).

However, since the tree is only an estimate of phylogenetic relationships, it was essential to verify the reliability of that estimate. The most common way to estimate the

reliability of a phylogenetic tree is by the bootstrap method (Felsenstein, 1985; Soltis and Soltis, 2003). It creates a new data set by sampling columns of characters randomly with replacement. The resulting new data set has the same length as the original, but some characters have been left out and others are duplicated. The method assumes that the characters evolve independently. After completion of the analysis the tree has numbers on every node, bootstrap percentages, which indicate the reliability of the cluster descending from that node. Higher percentages reflect increased reliability of each node. Based on the phylogenetic tree analysis of the BLASTp output (Figure 8), the pallilysin sequence was placed as most closely related to the laminin-binding sequence from *T. paraluiscuniculi.* Both sequences were well separated from the rest of the homologues as evident by the length of their common branch. Because of the high bootstrap percentage at the major node (94%), the sequences from *T. pallidum, T. paraluiscuniculi, T. phagedenis, T. pedis, T. denticola, T.viscentii,* and *T. medium* were clustered together with good confidence. In an analogous manner, so were the sequences from *T. brennaborense, T. saccharophilum, T. socranskii* subsp. *socranskii, T. socranskii* subsp. *paredis, T. succinifaciens, Treponema JC4, T. bryantii, T. lecithinolyticum,* and *T. maltophilum.* However, the individual grouping of the sequences within this major group was not so clear. Because of the low bootstrap percentages at their individual nodes (20-30%), the nodes could be collapsed, in which case, for example, *T. brennaborense* may still be clustered together with *T. saccharophilum* or it could be paired with similar confidence with *T. succinifacien,* the two *T.socranskii* subspecies, or the group of *Treponema JC4* and *T. bryantii.*

In summary, the phylogenetic analysis placed pallilysin with 94% confidence in a well defined group together with the sequences from *T. paraluiscuniculi, T. phagedenis, T. pedis, T. denticola, T.viscentii,* and *T. medium.* However, the long pallilysin branch indicated that even in this group pallilysin diverged significantly from the rest of the group members with the exception of *T. paraluiscuniculi.*

Tabulating the phylogenetic relationship between the sequences allowed mapping of certain motifs and sequence changes onto the tree more easily. Table 4, based on the Dayhoff distance Neighbour Joining Tree, shows the major clades of the tree separated by thicker lines. The particular clustering within each group is represented by thinner lines. The spirochetes, whose sequences were clustered in the bigger groups, were also grouped together in a previous 16S rRNA analysis by Whitman (2010). Therefore, the phylogenetic analysis of the pallilysin sequence homologues was congruent with the true species phylogeny. Thus, specific sequence changes in key motifs in the homologous sequences could reflect the nature of the species they belong to.

The sequence changes in the pallilysin laminin-binding and HEXXH motifs were traced on the phylogenetic tree and are presented in Table 4. As discussed earlier, the minimum pallilysin laminin binding region was previously mapped to three short amino acid sequences (Cameron *et al*., 2005). The $W^{127}I^{128}$ and the $T^{182}A^{183}I^{184}S^{185}$ motifs, presented in the table, form roughly two-thirds of the core laminin-binding region of pallilysin. The other four amino acids from the minimum laminin-binding region lie within a region of the protein that is too divergent to be accurately aligned, thus eliminating the possibility of accurate assessment of the conservation of this motif among the homologues.

Table 4. Phylogenetic relationships among pallilysin homologues.
The phylogenetic relationships among the pallilysin homologues were tabulated based on the Neighbour-joining Dayhoff distance tree with bootstrap analysis. The sequences are represented by their respective bacterial species with the pathogenic species and their association status highlighted in red. The sequence changes of the pallilysin laminin-binding (Cameron et al., 2005) and the HEXXH (Houston et al., 2011) motifs were mapped on the phylogenetic tree

Bacteria	**Pathogenic status and association**	**Laminin-binding motifs**		**HEXXH motif**
Treponema brennaborense	Pathogen: cattle (hoofs)	WY	GSVT	IDEIK
Treponema saccharophilum	Non-pathogenic: cattle (rumen)	WY	KSIE	LDEIR
Treponema socranskii subsp. *socranskii*	Pathogen: human (oral)	WY	KSIT	IDEVR
Treponema socranskii subsp. *socranskii*	Pathogen: human (oral)	WY	KSIT	IDEVR
Treponema succinifaciens	Non-pathogen: swine (intestine)	WY	QEIE	TEEIH
Treponema JC4	Non-pathogenic: cattle (rumen)	WY	SDIL	TDEIR
Treponema bryantii	Non-pathogenic: cattle (rumen)	WY	ADIM	TDEIK
Treponema lecithinolyticum	Pathogen: human (oral)	WV	KSMT	TDEIR
Treponema maltophilum	Pathogen: human (oral)	WT	RSMP	ADEIR
Treponema vincentii	Pathogen: human (oral)	WY	ASIS	IDEVR
Treponema medium	Pathogen: human (oral)	WY	ASIS	IDEVH
Treponema pallidum subsp. *pallidum*	Pathogen: human (systemic)	WI	TAIS	HEVIH
Treponema paraluiscuniculi cuniculi	Pathogen: rabbit (systemic)	WI	TAIS	HEVIH
Treponema phagedenis	Non-pathogen: human (skin/urogenital)	WY	ASIS	VDEIQ
Treponema pedis	Pathogen: cattle (hoofs)	WY	KSIP	VDEIQ
Treponema denticola	Pathogen: human (oral)	WF	KSIP	LDEIQ
Treponema caldaria	Non-pathogen: free-living thermophile	WF	LSVT	RDSIA
Treponema azotonutricium	Non-pathogen: termite (hindgut)	WY	ISVT	LDSIR
Treponema primitia	Non-pathogen: termite (hindgut)	WY	ISVT	LDSIR
Spirochaeta thermophila	Non-pathogen: free-living thermopile	WF	KTLP	ANTIQ
Spirochaeta africana	Non-pathogen: free-living haloalkaliphile	WY	QILP	LETIR

The tryptophane (W) of the WI sequence is part of the highly conserved GLW sequence as seen from the WebLogos (Figure 8). The GLW sequence was completely conserved among all treponemal homologues, and changed to GVW in the borrelia species. The leucine to valine point mutation is a conservative substitution since both leucine and valine are hydrophobic, non-polar amino acids of similar size. Therefore, the tryptophane (W) of the WI laminin-binding motif is fully conserved among treponemal, spirochaetal, and borrelial species. In addition, the directly preceding amino acids were also conserved providing a good alignment frame verifying the alignment. The

isoleucine (I) from the WI motif was only conserved in *T. pallidum* and *T. paraluiscuniculi* and is most commonly replaced by a tyrosine (Y) in the rest of the homologous sequences. A possible explanation could be found in the analysis of the amino acid frequencies in the multiple sequence alignments. Tyrosines were found at a much lower frequency in the pallilysin sequence than they were in the rest of the homologues.

An asparagine (N) directly preceded the other laminin-binding motif (TAIS) and was conserved in all homologous treponemal and spirochaetal sequences. The TAIS motif itself did not appear to be fully conserved, except in *T. pallidum* and *T. paraluiscuniculi.* However, in some of the homologues up to 3 conserved amino acids could be found in scrambled order. This laminin binding region (TAIS) is not sequence order specific, which makes the presence of some of the matching amino acids in some of the other homologues intriguing (Cameron *et al*., 2005). The highest degree of sequence conservation was concentrated in the treponemal oral pathogens. The *T. pallidum* TAIS motif appeared most closely related to the ASIS motif in *T. vincentii* and *T. medium,* both of which are human oral pathogens, as well as in *T. phagedenis,* which was originally isolated from a syphilitic chancre (Dashper *et al*., 2011; Trott *et al*., 2003; Wallace and Harris, 1967). In addition, the homologous sequences of these treponemes also exhibited the highest percent identities to pallilysin. Wild-type *T. phagedenis* lacks, or perhaps lost, the ability to attach to laminin, while *T. phagedenis* expressing pallilysin on its surface attached to laminin coated sides (Cameron et al., 2008a). Thus, perhaps the ASIS motif is distant enough from the TAIS motif in *T. pallidum* for it to not mediate laminin-binding. *T. pedis* and *T. denticola* are both pathogenic species and both have a KSIP motif. *T.*

pedis is a pathogen of cattle hoofs and *T. denticola* is a human oral pathogen (Dashper *et al*., 2011; Evans *et al*., 2009). While *T. denticola* was shown to attach to laminin (Haapasalo *et al*., 1991), the *T. denticola* homologue TDE0840 showed levels of laminin binding three times lower than those exhibited by pallilysin (Houston/Hof *et al*., unpublished observations). The species discussed so far all belonged to one major phylogenetic clade on the estimated tree. The next closest in sequence motif (KSIT) belonged to the two *T. socranskii* subspecies, both of which are implicated as human oral pathogens (Lee *et al*., 2006). They also belonged to the next major phylogenetic group. Another two human oral pathogens, *T. lecithinolyticum* and *T. maltophilum*, were found in this group with motifs, KSMT and RSMP, respectively (Park *et al*., 2002).

Finally, as apparent from both the WebLogos (Figure 8) and the phylogenetic table, the HEXXH motif itself is not to be conserved in its full sequence. However, the flanking regions showed good conservation (Figure 8), which provided a reliable alignment frame for the motif and validated the alignment. It is interesting to reiterate that the isoleucine (I), which is part of the HEVIH motif in pallilysin, although not essential for the metalloprotease function of the HEXXH motif, seemed to be almost fully conserved among treponema, spirochaeta, and borrelia species. In only very few cases, the isoleucine was replaced by a valine, which as previously discussed is a conservative substitution. The fact that the frequency of isoleucines in the pallilysin sequence was about half that in the rest of the homologous sequences, yet it was conserved in this motif, placed even greater importance on its conservation at that particular locale. The fact that this particular amino acid is found within the HEXXH motif, yet is not important for its functioning, but is so well conserved, again advocate for an additional ancestral

function of pallilysin that is conserved among the rest of the homologues. Interestingly, the glutamate (E) in the HEXXH motif essential for the protease role of pallilysin was most commonly changed to an aspartate (D), which however constitutes a conservative substitution.

Interestingly, the HEXXH counterpart motifs in the sequences from *T. phagedenis* and *T. pedis,* were identical. In addition, the laminin-binding motifs analyzed differed by only one amino acid. Taken together, these observations support the placement of the two sequences within the same group by the phylogenetic analysis. However, the two treponemes occupy not just different hosts (human vs. bovine), but also very different anatomical locations (genitalia vs. hoofs). Recent biochemical and molecular characterization of *T.phagedenis*-like spirochetes isolated from a bovine digital dermatitis lesion proposed that the description of *T. phagedenis* be expanded to include both human commensal and putative bovine pathogen (Wilson-welder *et al*., 2013). The latter could explain the high degree of sequence similarity between *T. phagedenis* and *T. pedis.*

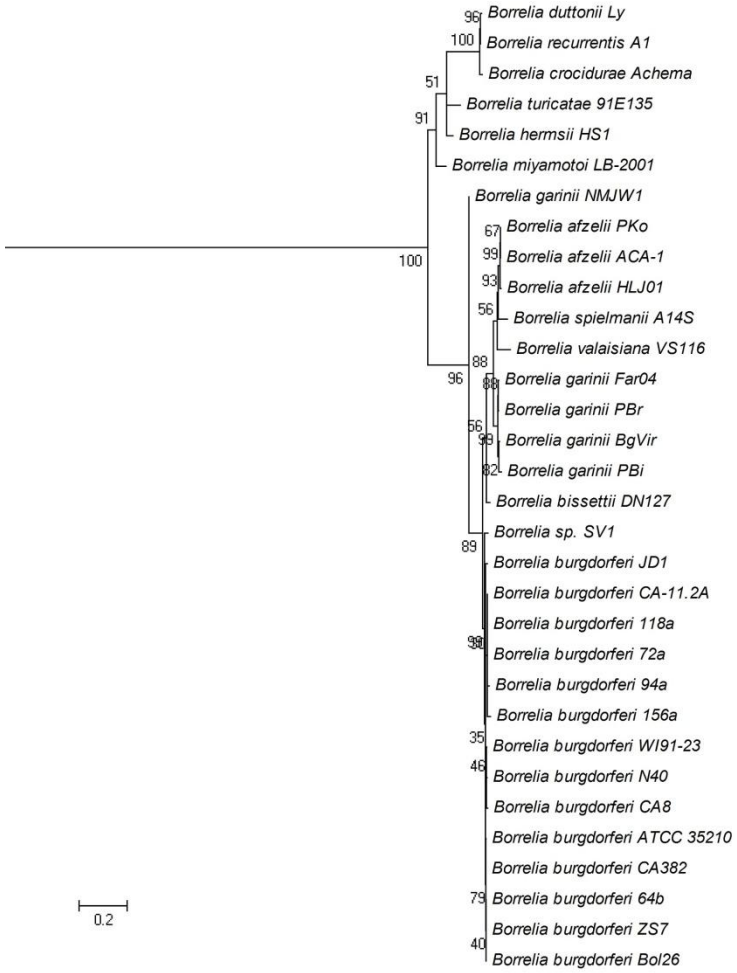
It is important to note that the function of a single protein, such as pallilysin, should not be equated to and does not exclusively explain the pathogenic profile of the species it belongs to; rather it provides another piece of the puzzle. Likewise, the degree of conservation of the pallilysin laminin-binding and the HEXXH motifs in the pallilysin homologues should not be solely and exclusively used to account for the nature of the species they belong to. Likely, a complex interplay of various factors, including multiple proteins, contributes to the pathogenic status of the bacteria. Therefore, the lack of a conserved HEXXH motif, for example, in some of the other pathogenic treponemes from the phylogenetic tree, should not be confused with lack of importance of this motif for

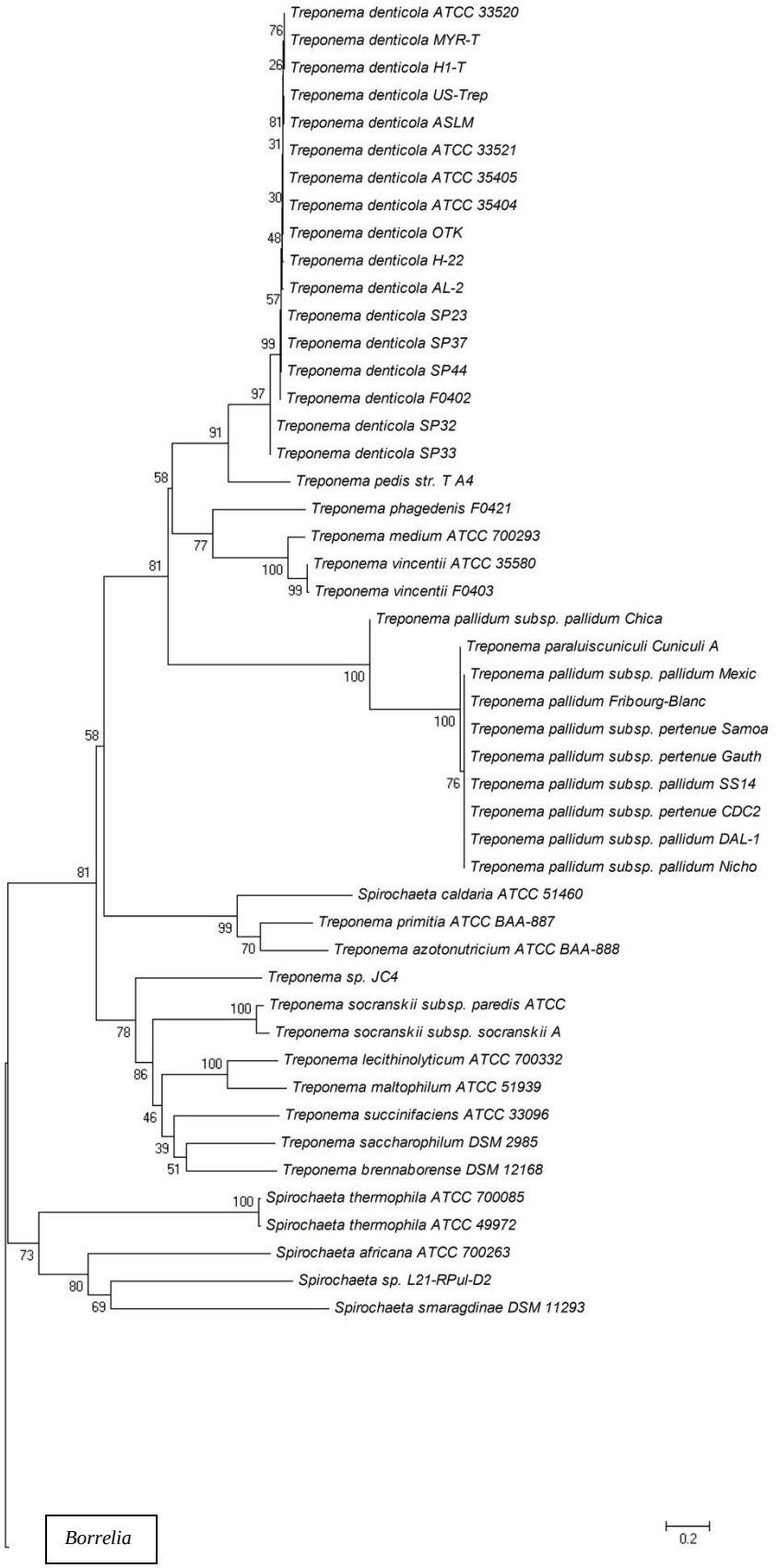
pathogenicity. Instead, it provides insight into the specific role of the HEXXH motif in pallilysin in comparison to the functions of the rest of the homologues. That in turn, provides further support for the role of pallilysin in establishing a disseminated infection by *T. pallidum* in contrast to the localized infections caused by the other pathogenic treponemes. However, further analysis of the pallilysin homologues through ECM binding and degradation assays is necessary to evaluate the results from the bioinformatics analyses.

Another phylogenetic analysis was performed based on the multiple sequence alignment from the HMMER search (Figure 9), which included additional spirochaeta and borrelia species not present in the BLASTp alignment. Two Neighbour joining trees were estimated: one was based on the same truncated region (not shown) and the other one contained the full length sequences of the homologues. The former resulted in a tree that placed some of the homologues present in the original BLASTp tree differently; however, the bootstrap evaluation produced low percentage values at the nodes. Therefore, the grouping of the sequences could not be accepted with confidence. It is likely that the newly introduced gaps in the alignment due to the additional sequences caused the difference in the arrangement of some of the homologues. In contrast, when the tree was estimated based on the full length multiple sequence alignment, the bootstrap re-analysis produced well defined groups of homologues similar to the original BLASTp tree arrangement. Homologous sequences from *T. pallidum* subsp. *pertenue*, not present in the BLASTp alignment, were included in the pallilysin clade, as expected. The homologous sequences from the borrelia species were placed together in a well defined group on a long branch signifying the further divergence from the treponema and

spirochaeta sequences. It is likely that the truncated region that contained the HEXXH motif and part of the laminin binding sequence in pallilysin was part of a domain that was well conserved among treponema, spirochaeta, and to a lesser extent among borrelia species. However, the flanking sequences of the truncated region exhibited much more pronounced divergence among the homologues and likely contributed to the separation of the sequences in well-defined clades on the phylogenetic tree.

Figure 9. Evolutionary relationships of 80 HMMER pallilysin homologues.
The evolutionary history was inferred using the Neighbor-Joining method. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the sequences analyzed. The percentage of replicate trees in which the associated sequences clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Dayhoff matrix based method and are in the units of the number of amino acid substitutions per site. All positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons (Pairwise deletion option). There were a total of 875 amino acids in the final dataset. Phylogenetic analyses were conducted in MEGA4 (Dayhoff *et al*., 1978; Gascuel and Steel, 2006a; Soltis and Soltis, 2003; Tamura *et al*., 2007).




Treponema denticola ATCC 33520
Treponema denticola MYR-T
Treponema denticola H1-T
Treponema denticola US-Trep
Treponema denticola ASLM
Treponema denticola ATCC 33521
Treponema denticola ATCC 35405
Treponema denticola ATCC 35404
Treponema denticola OTK
Treponema denticola H-22
Treponema denticola AL-2
Treponema denticola SP23
Treponema denticola SP37
Treponema denticola SP44
Treponema denticola F0402
Treponema denticola SP32
Treponema denticola SP33
Treponema pedis str. T A4
Treponema phagedenis F0421
Treponema medium ATCC 700293
Treponema vincentii ATCC 35580
Treponema vincentii F0403
Treponema pallidum subsp. pallidum Chica
Treponema paraluiscuniculi Cuniculi A
Treponema pallidum subsp. pallidum Mexic
Treponema pallidum Fribourg-Blanc
Treponema pallidum subsp. pertenue Samoa
Treponema pallidum subsp. pertenue Gauth
Treponema pallidum subsp. pallidum SS14
Treponema pallidum subsp. pertenue CDC2
Treponema pallidum subsp. pallidum DAL-1
Treponema pallidum subsp. pallidum Nicho
Spirochaeta caldaria ATCC 51460
Treponema primitia ATCC BAA-887
Treponema azotonutricium ATCC BAA-888
Treponema sp. JC4
Treponema socranskii subsp. paredis ATCC
Treponema socranskii subsp. socranskii A
Treponema lecithinolyticum ATCC 700332
Treponema maltophilum ATCC 51939
Treponema succinifaciens ATCC 33096
Treponema saccharophilum DSM 2985
Treponema brennaborense DSM 12168
Spirochaeta thermophila ATCC 700085
Spirochaeta thermophila ATCC 49972
Spirochaeta africana ATCC 700263
Spirochaeta sp. L21-RPul-D2
Spirochaeta smaragdinae DSM 11293
Borrelia
0.2

3.4 Conclusions

In order to gain further insight into the role of pallilysin in the dissemination of *T. pallidum*, a bioinformatics approach was undertaken. The pallilysin sequence was used as query to search the NCBI and UniProtKB protein databases to locate homologous sequences in an attempt to analyze the evolutionary conservation of the full length pallilysin sequence and that of important functional motifs. Both position-independent (BLASTp) and position-dependent (PSI-BLAST and HMMER) scoring matrices were used to search for homologues. The results included protein sequences not only from the treponema, but also from the spirochaeta and borrelia genera. Highest percent identity was observed with the homologue from *T. paraluiscuniculi.* The *T. phagedenis* homologue was next, followed by other sequences from treponeme species implicated as human oral pathogens. More distant homologues also included sequences from non-pathogenic treponema species, as well as sequences from borrelia and spirochaeta species. No homologues were detected from the fourth member of the *Spirochaetaceae* family, *Cristispira*, or from the other families in the order. Therefore, it can be concluded that the pallilysin sequence is evolutionarily conserved within the aforementioned genera in the *Spirochaetaceae* family. However, it is the degree of conservation of specific motifs in the homologues that might be integral to the role of the protein in pathogenesis.

Multiple sequence alignments of the homologous sequences allowed creation of sequence logos and phylogenetic trees. The former indicated the presence of a few very well conserved amino acids within the multiple sequence alignment in further support of the conservation of the pallilysin sequence. Upon a closer look at the degree of conservation of the HEXXH motifs, it became clear that the HEXXH motif, important for proteolysis, was only fully conserved in *T. pallidum* and *T. paraluiscuniculi*. However, an

isoleucine residue, part of the HEXXH motif in pallilysin (HEVIH), was almost fully conserved among the treponema, spirochaeta, and borrelia homologues. Therefore, while the HEXXH motif might have evolved in *T. pallidum* and *T. paraluiscuniculi* in accordance with the disseminated nature of their pathogenic profiles, part of the motif and the flanking sequences were very well conserved among the rest of the homologues. Therefore, the pallilysin homologues likely shared an alternate ancestral function. The fact that some of the homologues belong to free-living and non-pathogenic species in addition to the pathogenic ones indicates that the conserved ancestral function may not necessarily of pathogenic origin. In addition, the conservation of particular motifs and variation of others in the pallilysin sequence could have provided the necessary scaffold to facilitate the evolution and maintenance of a *de novo* motif, that of a metalloprotease. Furthermore, the two laminin-binding motifs investigated here were also conserved to varying degrees among treponema and spirochaeta species, more so in the human oral treponemes. It is yet to be tested if the homologues from all of these treponemes also bind laminin. The third laminin-binding motif, as mentioned earlier, was part of the highly variable region in the multiple sequence alignment and was left out of the phylogenetic analysis to prevent inaccurate interpretation of conservation.

Analysis of the amino acid frequencies in the pallilysin sequence compared to the rest of the sequences, represented by average amino acid frequencies, revealed that half of the amino acids were represented at the same frequencies, while the other half had either lower or higher frequencies in pallilysin. One of the most striking differences was the five times higher frequency of histidines in the pallilysin sequence, which could have facilitated the formation of the HEXXH motif. Analysis of the placement and

conservation of the other histidines in the pallilysin sequence could provide further insight into the evolutionary history and function of the protein.

Finally, the phylogenetic analysis of the homologues allowed grouping the more closely related sequences together and mapping amino acid changes in the laminin-binding and HEXXH motifs. As expected, homologues placed closer together on the tree had fewer amino acid substitutions in these motifs. The degree of conservation among the homologues supported the relatedness among the proteins, yet provided evidence of perhaps enough divergences to explain the role of pallilysin in dissemination consistent with a pathogen able to cause a systemic disseminated infection. Therefore, looking at pallilysin in the context of its homologues in a phylogenetic tree with mapped sequence changes provided further support for a role in the pathogenesis and wide-spread dissemination of *T. pallidum.*

Chapter 4: Characterizing the Dissemination Capacity of *Treponema phagedenis* heterologously expressing pallilysin

4.1 Contributions to the data

My contributions to the data presented in this chapter include development of the artificial endothelial barrier, further development of the dissemination assay methodology based on earlier studies in our laboratory, and the execution of the dissemination assays and the quantification of percent traversal through dark-field microscopy and real time quantitative PCR. Charmaine Wetherell assisted in the execution of the assays, and had originally developed the real time quantitative PCR methodology for the project at the University of Victoria, Canada.

4.2 Introduction

T. pallidum is able to spread to every organ in the human body via the blood circulation causing a persistent disseminated infection (Lafond and Lukehart, 2006). Hematogenous spread of the pathogen requires that *T. pallidum* cross the vascular endothelial barrier and the underlining basement membrane (Haake and Lovett, 1994a). Indeed, virulent *T. pallidum* can adhere to basement membranes, to a variety of ECM components, and to cultured endothelial cells, as well as penetrate their intercellular junctions (Fitzgerald *et al*., 1984; Lovett *et al*., 1988; Riley *et al*., 1992; Thomas *et al*., 1989). Interference with the adherence of the viable pathogens to cell cultures by immune sera suggested that attachment to host cells may be mediated by specific *T. pallidum* surface exposed antigens (Fitzgerald *et al*., 1977a; van der Sluis *et al*., 1987).

Pallilysin was first identified as a putative outer membrane laminin-binding adhesin (Cameron, 2003). Laminin is a major component of basement membranes, which implicated the protein in having a role in the dissemination of the pathogen. Later, attachment studies showed pallilysin was also able to specifically bind to fibrinogen, a major component of the coagulation cascade (Houston *et al*., 2011). Pallilysin was then definitively shown to be surface exposed in *T. pallidum* (Houston *et al*., 2012). Furthermore, a novel function was described for the protein: a metalloprotease capable of degrading laminin, fibrinogen, and fibrin clots (Houston *et al*., 2012). The latter serve to contain bacterial pathogens from spreading. Finally, heterologous expression of pallilysin in *T. phagedenis* conferred upon this non-adhesive and non-invasive *treponeme* the capacity to attach to laminin and degrade fibrin clots (Cameron et al., 2008a; Houston et al., 2012). Therefore, to build upon the results accumulated so far in support of a role of pallilysin in the dissemination of *T. pallidum*, we set out to find out whether heterologous expression of pallilysin in *T. phagedenis* also conferred upon this bacterium dissemination capacity; that is, ability to penetrate a vascular endothelial barrier. Unlike *T. pallidum, T. phagedenis* can be cultured in growth medium, which makes laboratory manipulations much more feasible. Furthermore, the fact that *T. phagedenis* is more amenable to genetic manipulations in the first place allowed transformation of wildtype *T. phagedenis* with both the empty shuttle vector, pKMR, and the shuttle vector containing the full open reading frame of pallilysin (Cameron et al., 2008a). As described in the bioinformatics chapter, a homologue of pallilysin does exists in *T. phagedenis*; however, the percent identity is 42%, and the HEXXH motif, which is integral for the metalloprotease function of pallilysin, is not conserved in the *T. phagedenis* homologue.

4.3 Dissemination Assay Design

Hematogenous spread of *T. pallidum* requires that the pathogenic bacteria cross the endothelial lining of blood vessels. Therefore, in order to traverse from the lumen to the surrounding tissues, the bacteria have to first penetrate the tight junction-containing endothelial monolayer and the underlying basement membrane. To test whether heterologous expression of pallilysin in *T. phagedenis* conferred dissemination capacity on this otherwise non-adhesive and non-invasive bacterium, a transwell dissemination assay was developed. The assay was based on a technique initially developed for studying macromolecular transport and leukocyte transcytosis, and later on adapted for spirochete dissemination studies (Haake and Lovett, 1994a). The latter include studies focused on the interactions with, and penetration of, endothelial or epithelial layers by *T. pallidum* (Lovett *et al*., 1988; Riviere *et al*., 1989; Thomas *et al*., 1989), *T. denticola* and other oral spirochetes (Lux *et al*., 2001; Peters *et al*., 1999; Riviere *et al*., 1991), *B. burgdorferi* (Comstock and Thomas, 1989; Grab *et al*., 2009), and *Leptospira* species (Thomas and Higbie, 1990). A similar approach was used to study the interaction of virulent *Trypanosoma* with a model of the blood-brain-barrier (Grab *et al*., 2013). All of the studies mentioned above shared a common basic design: the dissemination chamber was composed of two compartments separated by a membrane that allowed for a serum chemotactic gradient to be formed between the top and the bottom compartment. The membrane was made out of mouse abdominal walls, or either epithelial or endothelial cells. When epithelial or endothelial cell layers were used, they were grown on permeable polycarbonate or polyester filters with pores of different densities and diameters. In some

cases, the permeable filters were coated with a basement membrane component, such as fibronectin, or gelatin, upon which the endothelial or epithelial cells was seeded. Finally, the microorganisms of interest were added to the top compartments, or otherwise called inserts, and adherence to and penetration of the monolayers were investigated. The presence of an upper and a lower compartment facilitated the assessment of intercellular junction integrity by means of transendothelial or transepithelial electrical resistance measurements and allowed quantification of the traversal of the microorganisms (Haake and Lovett, 1994a).

In the studies of the hematogenous spread of spirochaetes, HUVEC cells were commonly used, and while some researchers preferred to isolate them fresh from the human umbilical vein, they are also available commercially in frozen stocks. Provided that a low passage number (less than 5 passages) was used, the HUVECs would maintain tight junctions and thus their monolayers would be suitable for bacterial dissemination studies (Haake and Lovett, 1994a; Man *et al*., 2008). As discussed earlier, the integrity of the monolayers was assessed before addition of the microorganisms and at the end of the incubation period by measuring the trans-endothelial electrical resistance (TEER). The TEER varies with different cell lines reflecting their respective tight junctions and the demand of their underlining tissues. Decreased resistance resulted in increased vascular permeability (Claude, 1978). Thus, the endothelial cells that make up the blood-brain-barrier exhibit very high TEER readings of 1500–2000 $OhmsXcm^2$ consistent with their role to maintain homeostasis of the brain (Huber *et al*., 2001). In comparison, confluent monolayers of HUVECs were reported to exhibit readings that vary from less than 12 $Ohms*cm^2$ (Grab *et al*., 2005) to 21 $OhmsXcm^2$ (Amerongen *et al*., 1998) and reach as

high as 74 OhmsXcm^2 at passage 3 with a drop to 33 OhmsXcm^2 at passages 5-7 (Man *et al*., 2008). Therefore, it was suggested that the passage number and the culture conditions greatly affected the tight junction protein expression in HUVECs (Haake and Lovett, 1994a; Man *et al*., 2008). Therefore, in addition to the seeding density and growth rate, the passage number factors in the rate at which the monolayer becomes confluent. The acceptable threshold HUVEC TEER value for spirochete dissemination studies has been previously established as 10 OhmsXcm^2 (Haake and Lovett, 1994a).

4.3.1 Artificial endothelial barrier

A few different dissemination chamber designs were tested in the current dissemination study, including transwell plates with polycarbonate (PC) and polyester (PET) membranes, with pore sizes varying from 3 μm to 8 μm, as well as different coating agents, such as fibronectin, gelatin (0.1, 1, and 2%), and Matrigel. The HUVECs could not proliferate on the PC filters, but were able to attach and proliferate on the PET filters. The coating agents had a positive effect on the growth of the HUVECs on the permeable inserts with the best results obtained with the BD Matrigel Matrix Basement Membrane coating. Thus, the transwell plate chosen was a 24-well plate with 12 transwells with 8 μm pore size PET filters pre-coated with a thin layer of BD Matrigel Matrix Basement Membrane. Since, the Matrigel Matrix is a reconstituted basement membrane that contains the same components as *in vivo* basement membranes, that arrangement provided the ideal experimental conditions to create an artificial endothelial barrier (BD Biosciences, 2011). In addition, laminin, a major component of basement membranes, and a proteolytic substrate for pallilysin, also constitutes the biggest

component in the Matrigel Matrix Basement Membrane. Seeding the HUVECs on top of the Matrigel Matrix Basement Membrane essentially reproduced the lining of the blood vessels that virulent *T. pallidum* must have a mechanism to cross for hematogenous spread to occur. Therefore, we could test whether pallilysin was directly involved in the penetration of the endothelium by assaying whether heterologous expression of pallilysin in *T. phagedenis* allowed that bacterium to traverse the artificial endothelial barrier. In order to provide a chemotactic stimulus for the treponemes to invade the HUVEC monolayer, the bottom compartment of the dissemination chamber was supplied with twice the normal rabbit serum concentration found in the top compartment, where the bacteria were added. Therefore, a chemotactic gradient was formed across the transwell. Thus, provided that monolayer confluence in control transwells, where no treponemes were added, was maintained for the duration of the assay, transendothelial traversal by *T. phagedenis* heterologously expressing pallilsyin would directly implicate pallilysin in the process of endothelial barrier invasion and hematogenous dissemination of *T. pallidum.* Traversal was assessed in comparison to that of wildtype and shuttle-vector transformed *T. phagedenis* as negative controls. In addition, the integrity of the HUVEC monolayers in the presence of either strain of *T. phagedenis* also was assessed.

4.3.2 Optimizations of the Dissemination Assay

Previous studies from the late 1980's on the penetration of endothelial, including HUVEC, monolayers by *T. pallidum* and *B. burgdorferi*, with *T. phagedenis* as a negative control, were performed in 5% CO_2 and 95% air (Comstock and Thomas, 1989; Lovett *et al*., 1988). A decade later, a study on the adherence and penetration of endothelial layers

by oral treponemes including *T. denticola*, and also using *T. phagedenis* as a negative control, was still performed under the same atmospheric conditions (Peters *et al*., 1999). It is surprising that the spirochetes maintained viability under these conditions keeping in mind that *T. pallidum* is microaerophilic, and *T. phadenis* and *T. denticola* are anaerobic. Perhaps the short duration of the experiments (6, 4 and 12 hours, respectively) contributed to the success of the assays. In 1994, a review was published outlining the methodology behind dissemination studies with spirochetes and suggesting incubation of the endothelial monolayer and the bacteria under the microaerophilic conditions of 5% O_2, 5% CO_2, 90% N_2 (Haake and Lovett, 1994a). Such conditions should also maintain the TEER of the endothelial monolayer, since the 5% O_2 and 5% CO_2 translate to partial pressures of oxygen and carbon dioxide at levels very close to those found in the venous circulation (Carreau *et al*., 2011; Haake and Lovett, 1994b). In a 2001 study the penetration of epithelial tissue layers of immortalized keratinocytes by *T. denticola*, as well as *T. pallidum* (positive control) and *T. phagedenis* (negative control),anaerobic conditions were used (85% N_2, 10% H_2, 5% CO_2) (Lux *et al*., 2001). The authors pointed out that the reason for maintaining such conditions was because of rapid loss of motility of *T. denticola* in aerobic conditions. However, anaerobic conditions are not optimal for the survival of mammalian monolayers, so pyruvate and uridine were added to the growth medium of the epithelial cells to maintain tissue viability for 2 to 3 days under anaerobic conditions. In addition, the cells used were immortalized and it is possible that that also helped them withstand the lack of oxygen for a longer period.

Based on the information from these studies, a few different conditions were attempted for the current assay. The HUVECs were initially grown in a 5% CO_2

incubator (95% air), while *T. phagedenis* cultures were maintained in an anaerobic chamber (2% H_2, 98% N_2). *T. phagedenis* is considered a genital skin commensal of the human flora; therefore, such association requires a mechanism of co-existence of the spirochetes and the human cells. However, *T. phagedenis* has been continually passaged in culture medium under anaerobic conditions since its original isolation. When growth under microaerophilic conditions was attempted, after 24 hours in our hands, *T. phagedenis* had completely lost motility. Therefore, the only feasible option left was to design the dissemination assay under anaerobic conditions. That, however, posed a problem for the HUVECs since the cell line was a primary cell line that required oxygen to maintain viability and, as an extension, tight junctions in the monolayer. Therefore, a multi-prong approach was utilized to bridge these two diametrically opposed growth conditions in a laboratory setting.

Firstly, *T. phagedenis* was maintained in liquid culture in TYGVS medium prepared in the laboratory, while the HUVECs were cultured in Lonza's proprietary EGM-2 Bulletkit medium. Viability of either was drastically reduced when grown in the opposite medium; therefore, a mixture of both media was prepared to support the survival and growth of both HUVECs and treponemes. The best results for co-incubation under anaerobic conditions were obtained with a 1:4 mix of TYGVS and EGM-2 Bulletkit media, respectively.

Since anaerobic culture conditions mimic a phenotype caused by a deficiency of electron transport chain proteins, eukaryotic cells transition into pyruvate and pyrimidine auxotrophs (Bodnar *et al*., 1995). The addition of 55 mg/L of pyruvate and 110 mg/liter uridine to the growth medium allowed immortalized epithelial tissue layers to remain

healthy for 2-3 days under anaerobic condition (Lux *et al*., 2001). Pyruvate was already available in sufficient concentration in the HUVEC medium used in the current study. In our hands, however, addition of uridine at the suggested concentration, as well as higher and lower concentrations, did not have an obvious effect on the morphology of the HUVECs. The difference in the effect of uridine in the HUVEC and the immortalized epithelial tissue layers could have been due to the different tissue origins and the fact that the HUVECs were a primary cell line. Unlike the addition of uridine, increasing the fetal bovine serum concentration in the EGM-2 Bulletkit medium from the original 2% to 12% resulted in higher TEER readings indicative of more confluent monolayers.

Further optimization of the growth medium was required to maintain physiological pH in anaerobic conditions. The EGM-2 Bulletkit medium, which is the major component of the mixed medium, is bicarbonate based. So, its pH thus depends on the presence and percentage of carbon dioxide in the surrounding atmosphere. It is adapted to maintain physiological pH at 5% CO_2 in the atmosphere. However, the anaerobic chamber where the dissemination assay was conducted only had hydrogen and nitrogen gas continually pumped in with a catalyst responsible for binding any available oxygen molecules to hydrogen. Therefore, the lack of carbon dioxide resulted in a steep increase in the mixed medium pH reaching values of over pH 8. To maintain the pH within the physiological range, 25mM HEPES was added to the medium.

In the process of optimization of the growth medium composition, an interesting phenomenon was observed. In a few of the wells where HUVECs were grown under anaerobic conditions, without the buffering capacity of HEPES, their cell morphology had changed. Confluent cells that form a monolayer exhibit cobblestone morphology,

while these ones were in various stages of differentiation into capillary tubes. Images of a confluent layer of HUVECs and differentiated HUVECs grown on Matrigel are shown in Figure 10. The characteristic cobble stone morphology and capillary tube structures, respectively, can be observed.

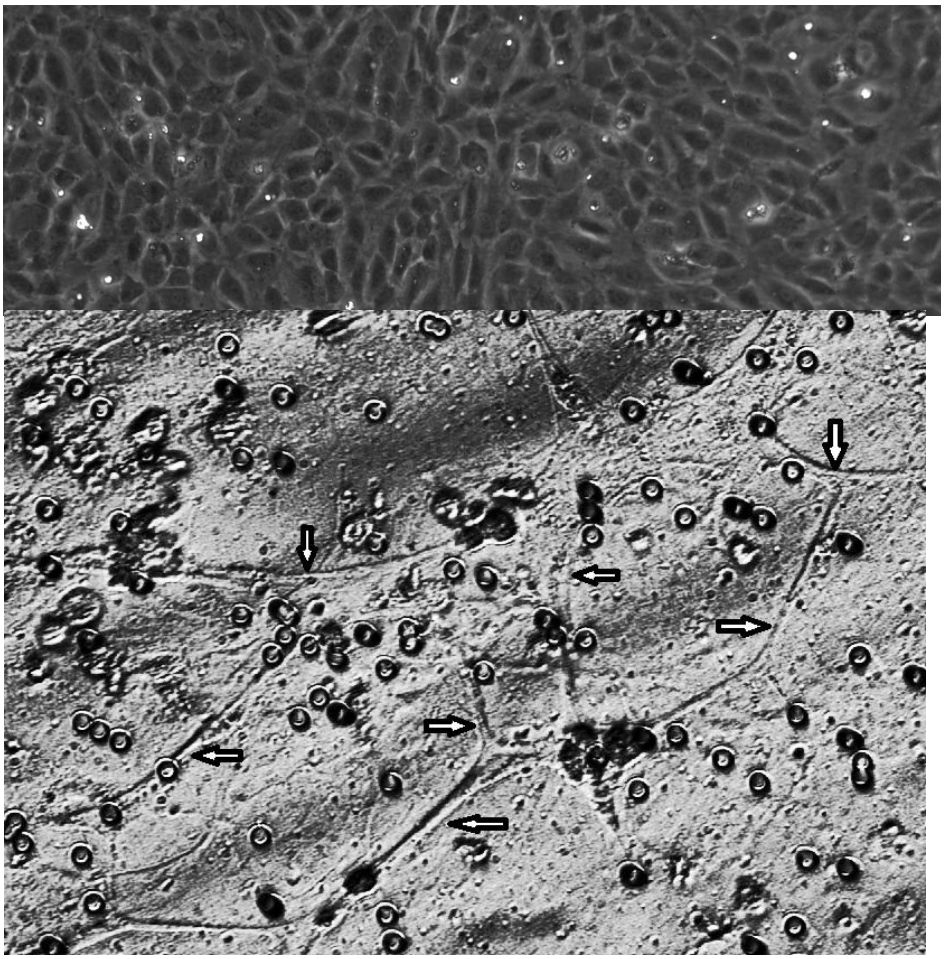


Figure 10. Microscope images of different HUVEC morphologies. Typical cobble stone morphology of a confluent layer of Human Umbillical Vein Endothelial Cells (HUVEC) grown on tissue culture treated polystyrene flasks (up) and HUVECs differentiated into a network of capillary tubes on Matrigel Matrix basement membrane coated permeable inserts (down). The HUVEC tubes are shown with arrows. The circular structures represent the pores of the permeable inserts upon which the endothelial cells are seeded.

Interestingly, tube formation was also observed in tissue culture flasks in the 5% CO_2 incubator left for about 2 weeks without exchanging the growth medium, which was consistent with the literature (Walker, 2009). The occurrence of such a phenomenon during the dissemination assay was one of the biggest concerns with the current design since usually HUVEC cells grown on Matrigel matrix tend to form a network of tube-like structures as part of the angiogenesis process. Such morphology however does not provide a good barrier for the treponemes in the dissemination assay. On one hand, the basement membranes underlying the endothelial layer have been implicated in the regulation of endothelial cell behaviour with laminin substratum promoting endothelial cell differentiation into capillary-like structures (Grant *et al*., 1990). When endothelial cells, such as HUVECs, are grown on Matrigel matrix, a reconstituted basement membrane rich in laminin, tube formation generally occurs within 18 hours. On the other hand, the lack of oxygen additionally contributes to tube formation. Neo-vascularization plays a crucial role in several pathological conditions, including tumor growth (Walker, 2009), and also in the process of blood vessel development in the mammalian embryo (Jin *et al*., 2012). In both cases, a hypoxic environment (1% O_2) stimulates the switch to an angiogenic phenotype. Under hypoxia, the transcriptional factor complex Hypoxia Inducible Factor (HIF) is activated and HIF in turn induces the secretion of growth factors, including the vascular endothelial growth factor (VEGF), which stimulate endothelial cells to differentiate into capillary-like structures (Jin *et al*., 2012; Walker, 2009). In a study performed to investigate the interplay of hypoxia and growth factors in the culture medium and their impact on capillary morphogenesis, HUVECs were seeded on growth factor-reduced Matrigel in the presence or absence of exogenously added

growth factors and exposed to either hypoxia (1% O_2) or normoxia (21% O_2) (Calvani *et al*., 2006). In the presence of growth factors, growth on Matrigel under either hypoxia or normoxia resulted in tube formation, while only hypoxia induced angiogenesis in the absence of growth factors. To prevent tube formation in the current study, three measures were taken to ensure the proliferation of the HUVECs and the integrity of the monolayer. Firstly, the pH was stabilized with HEPES as it was the pH fluctuation under anaerobic conditions that first resulted in angiogenesis in our hands. Second, since we deemed it unreasonable to deplete the culture medium of growth factors for the long duration of the experiment, we chose to instead use growth-factor reduced Matrigel, which not only contained reduced levels of growth factors but also provided a more highly defined basement preparation than regular Matrigel. The latter is essential since Matrigel is obtained from the Engelbreth-Holm-Swarm (EHS) mouse sarcoma, a tumor rich in extracellular matrix proteins and naturally occurring growth factors (BD Biosciences, 2011). In a study by Matsubara *et al*., growth factor-reduced Matrigel did not induce differentiation of uterine artery endothelial cells in the presence of estrogen, while complete Matrigel did (Matsubara *et al*., 2008). Finally, the thin gel culture method was used rather than the thick or 3D method. The thin Matrigel gel provides a substratum for attachment and proliferation of endothelial cells. It acts as a protein layer instead of a protein matrix and can thus support the propagation of primary cells, such as HUVECs (BD Biosciences, 2011). The dissemination chambers used in the current study were pre-coated with approximately 10µm thick Matrigel for use with the thin gel culture method.

After the substratum, the culture medium, and the atmosphere were defined, the last two variables left were the temperature and duration of incubation of the treponemes

with the HUVECs. When cultured separately, both HUVECs and *T. phagedenis* were grown at 37°C. However, it was quickly established that incubation of the HUVECs at 37°C in anaerobic conditions resulted in a rapid drop of TEER readings that was, in this case, independent of the presence or absence of the treponemes (Figure 11).

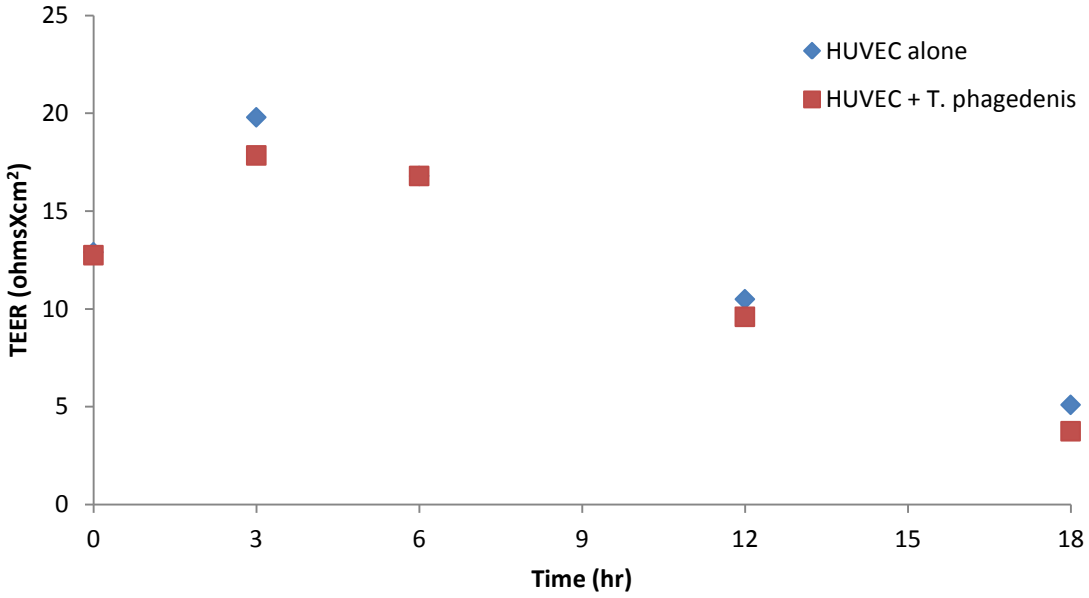


Figure 11. Failure to maintain HUVEC monolayer integrity past 12 hours at 37°C in anaerobic conditions. HUVECs were seeded on permeable Matrigel coated transwell inserts and incubated overnight at 37°C in the 5% CO_2 incubator to reach confluence. Inserts with HUVECs alone and HUVECs and *T. phagedenis* ($3.5x10^7$ pKMR/*T. phagedenis* per insert) were then transferred to anaerobic conditions at 37°C. TEER measurements (ohmsXcm^2) represent the level of monolayer integrity at different time points over a period of 18 hours. Measurements were performed on the same wells at each time point throughout the duration of the experiment.

It is possible that at 37°C the HUVECs are at optimal metabolic capacity and quickly use up any residual oxygen in their medium and the Matrigel after being transferred to the anaerobic chamber. Furthermore, from a kinetic point of view, the diffusion of oxygen at higher temperatures is faster than that at lower temperatures.

Therefore, oxygen diffusion out of the medium and into the anaerobic atmosphere of the chamber would occur faster at 37°C and the oxygen would be bound to hydrogen at the chamber catalyst. Finally, viability and proliferation rely heavily on the availability of energy, the majority of which is normally provided by glycolytic pathways that involves the electron transport chain with oxygen as the final electron acceptor. Lack of oxygen forces the cells to rely only on glycolysis to obtain energy, which yields significantly fewer ATP molecules that cannot meet the demands of an otherwise actively functioning cellular machinery. By dropping the incubation temperature, the residual oxygen diffuses more slowly and is thus available to the cells for longer. Furthermore, from an energetic point of view, lower temperatures would slow down cellular processes and could potentially allow the cells to withstand the lack of oxygen for a longer period. Decreasing the temperature down to 25°C indeed prolonged the integrity of the HUVEC monolayer as measured by TEER for 30 to 48 hours (Figure 12). A decrease in TEER was still observed at 25°C; however, it was much slower and more gradual in contrast to the sudden drop in TEER at 37°C between 12 and 18 hours. In fact, an increase in the TEER readings was recorded at 30 hours. Since it was crucial for the purposes of the dissemination assay to maintain a confluent monolayer as long as possible for the duration of the experiment, incubation at 25°C was chosen for the assay.

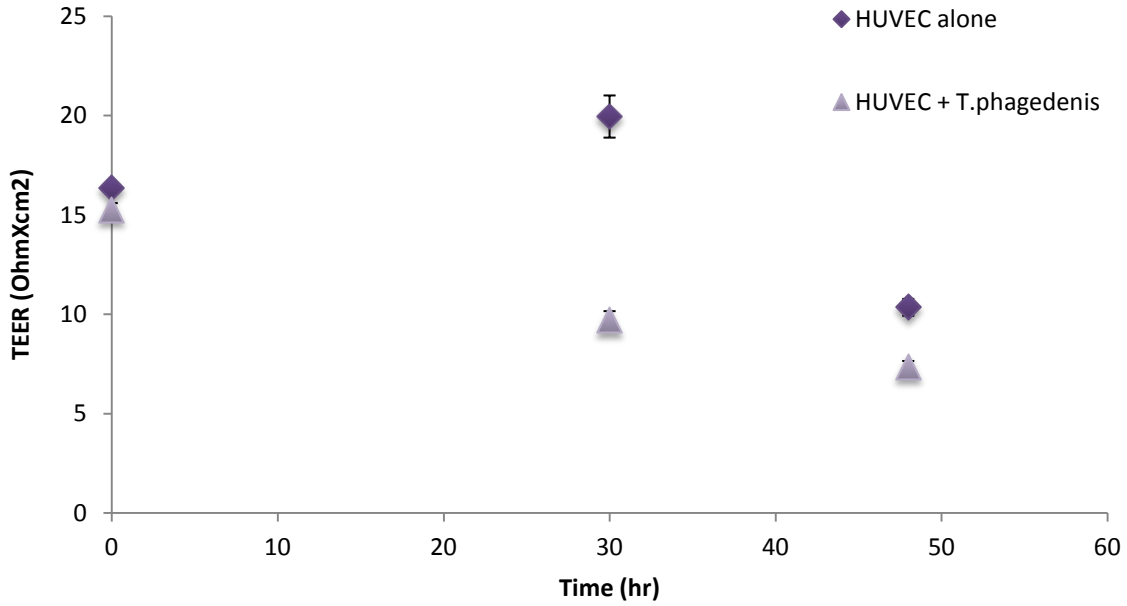


Figure 12. Incubation at 25°C allows extended maintenance of the HUVEC monolayer integrity under anaerobic conditions. TEER measurements of HUVEC monolayer alone and HUVEC monolayer with *T. phagedenis* ($3.5x10^7$ pKMR/*T. phagedenis* per insert) were taken at 0, 30, and 48 hours to assess monolayer integrity. After an initial increase in TEER at 30 hours, the readings for the HUVECs alone samples had decreased at 48 hours, but remained above the confluence threshold of 10 ohmsXcm^2. Addition of pKMR/ *T. phagedenis* appeared to further decrease the TEER readings over the incubation period.

Interestingly, after incubation at 37°C for 48 hours the treponemes appeared overgrown and were clumping in structures resembling mats. As a result, their motility was more similar to sideways twitching than the typical fast complete revolutions around their axes. In contrast, incubation at 25°C resulted in highly motile organisms that clumped significantly less. Therefore, the lower temperature slowed down the growth of the treponemes concurrently preserving nutrients in the medium for a longer period. The decrease in incubation temperature, albeit also beneficial for the extended viability of the HUVEC monolayer, raised the possibility of lower activity of the pallilysin protease on the surface of *T. phagedenis.* In order to verify that pallilysin would achieve sufficient

catalytic rate in the conditions of the dissemination assay, an SDS-PAGE fibrinogen degradation assay was carried out as described by Houston *et al*, 2012. Wildtype and active site mutant (HAXXH E199A) pallilysin were pre-incubated at 37^{0}C for 21h prior to addition of fibrinogen. The pallilysin pre-incubation period also reflected the growth period of the treponemes before plating for the dissemination assay. This pre-incubation step was included because previous studies in our laboratory conducted by Dr. Simon Houston established that full-length pallilysin was activated by intermolecular autocatalysis and exhibited a characteristic proteolytic activation lag period that was shortened by pre-incubation at 37°C for 21-22 hours (Houston *et al*., 2011, 2012). Incubation of pre-activated pallilysin with fibrinogen, as well as of fibrinogen alone, was then conducted at either 37^{0}C or 25^{0}C for 48 hours (Figure 13). Samples were taken at 0, 24 and 48 hours and run on an SDS-PAGE gel to detect degradation of the fibrinogen α-, β-, and γ-chains as a measure of the level of proteolytic activity of pallilysin. Similar to previous results by Houston *et al*. (2011, 2012), by 24 hours wild type pallilysin incubated with fibrinogen at 37^{0}C was able to degrade both the α- and β-chains of fibrinogen, followed by further degradation of the γ-chain by 48 hours. In contrast, when wild type pallilysin was incubated with fibrinogen at 25^{0}C, by 24 hours only the α-chain was degraded, and by 48 hours both the α- and the β-chains were degraded. Therefore, it appears that the activity of pallilysin is decreased at 25°C.

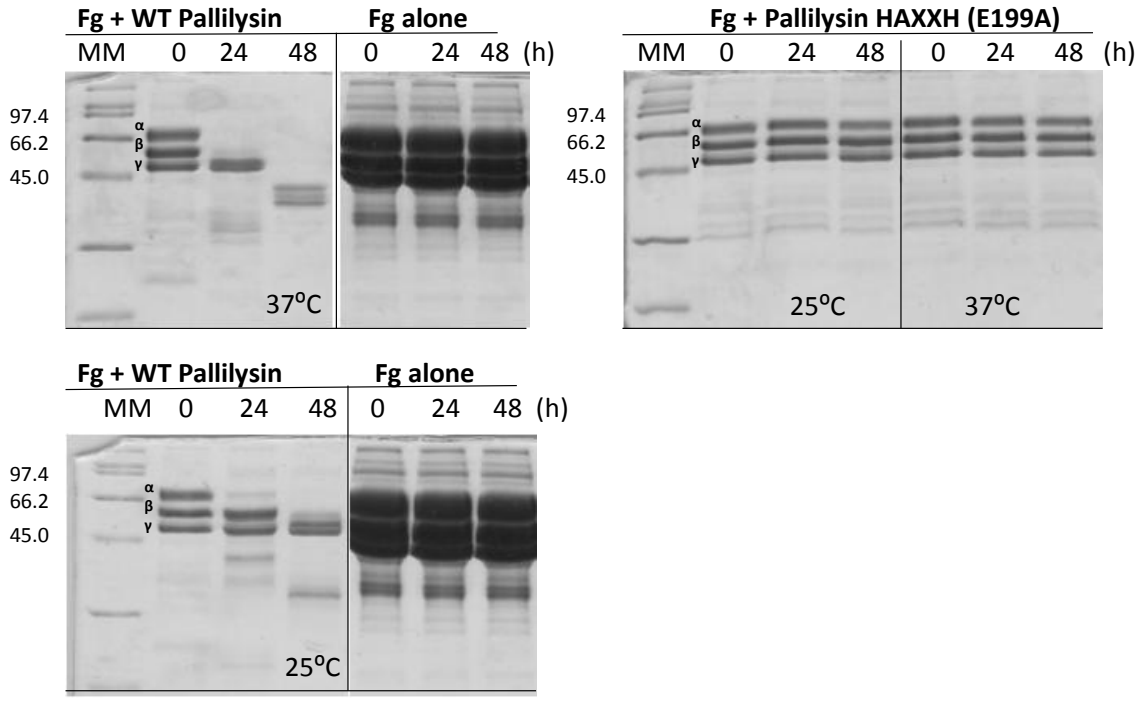


Figure 13. Catalytic activity of pallilysin is slowed down at 25°C. Wild-type (WT) pallilysin subjected to 21 h pre-incubation at 37°C in the absence of substrate was incubated with fibrinogen at either 37°C or 25°C for 48 h. Samples were removed at the times indicated and analyzed by SDS-PAGE to detect degradation of the fibrinogen α-, β-, and γ-chains. Negative controls included fibrinogen incubated in buffer alone (Fg only) and fibrinogen incubated in the presence of the HAXXH active site mutant pallilysin, which had also been pre-incubated at 37°C for 21 h.

The HAXXH active site mutant pallilysin was unable to degrade fibrinogen over 48 hours of incubation at either temperature. Fibrinogen alone was also stable for the duration of the assay. The fibrinogen only bands on the SDS-PAGE gel appear much thicker than the rest due to a concentration error. However, that did not undermine the validity of the conclusion since the right amount of fibrinogen was incubated with wild

type and HAXXH pallilysin, and no obvious degradation was detected by the HAXXH mutant. Therefore, fibrinogen itself remained stable for the duration of the experiment.

In conclusion, the necessity to incubate the treponemes and the HUVECs at a lower temperature (25°C) to maintain the HUVEC monolayer integrity for longer than 12 to 18 hours now posed a challenge for the activity of pallilysin. In order to ensure sufficient activity of the protease was achieved, the dissemination assay was designed to run for 48 hours. It should be noted that the pallilysin activity assay was performed on a recombinantly expressed protein; therefore, it is an estimate of the activity of pallilysin on the surface of *T. phagedenis.*

In summary, in order to find a solution to the challenging scenario of co-incubating the HUVECs and the treponemes, the dissemination assay was carried out under anaerobic conditions with a special culture media mix; the thin Matrigel culture method was utilized; the pH was controlled for and the incubation temperature and assay duration were adjusted accordingly.

4.3.3 Assessment of Dissemination Capacity of *T. phagedenis* Heterologously Expressing Pallilysin

The ability to invade an endothelial monolayer directly correlates with spirochete pathogenicity (Haake and Lovett, 1994a). Although the morphology and motility of pathogenic and non-pathogenic spirochete members are similar, their invasion capacity is strikingly different. Heat-inactivation renders spirochete pathogens non-motile and as a result they lose their invasion capacity. However, while motility is necessary, it is not sufficient for invasion of endothelial monolayers. The motile non-pathogenic *T. phagedenis* biotype Reiter penetrated endothelial and epithelial monolayers, and mouse

abdominal walls, only at about 10% of the rate of virulent *T. pallidum* (Comstock and Thomas, 1989; Lovett *et al*., 1988; Lux *et al*., 2001; Riviere *et al*., 1989). Similar results were obtained with the avirulent culture-attenuated strain B31 of *B. burgdorferi* and the virulent strain 297 of *B. burgdorferi* (Haake and Lovett, 1994a). Likewise, the non-pathogen *Leptospira biflexa* invades endothelial monolayers at a rate that is less than 10% the invasion rate of the virulent *Leptospira interrogans* (Thomas and Higbie, 1990).

The aim of the current study was to determine whether heterologous expression of pallilysin in the non-invasive *T. phagedenis* biotype Kazan conferred upon this bacterium dissemination capacity across the artificial endothelial barrier. As controls we used both wildtype *T. phagedenis* and *T. phagedenis* transformed with the empty shuttle vector (pKMR). Invasion capability for each strain were determined by quantifying the number of bacteria that crossed the artificial endothelial barrier relative to the total number of bacteria present at the end of the assay period. Quantification was carried out both by dark-field microscopy counts and by real-time quantitative PCR. Depending on the optics of the dark-field microscope, a 10µL sample under a $1cm^2$ coverslip containing a spirochete concentration of 10^6/ml results in roughly one organism per high-power field (Haake and Lovett, 1994a). In our hands, detection was variable in a 12µL sample of 10^5 treponemes/ml. Therefore, it was difficult to enumerate spirochetes accurately at concentrations below $l0^5$/ml and some of our samples had lower concentrations. In addition, *T. phagedenis* is prone to clumping, which made enumeration by dark-field microscopy challenging. For these reasons, QPCR methodologies were used to complement the dark-field microscopy quantification of bacterial traversal across the artificial endothelial barrier.

Comstock and Thomas (1989) reported endothelial monolayer traversal percentages of 9.60%, 5.80%, and 0.50% for *B. burgdorferi, T. pallidum*, and *T. phagedenis* over 4 hours of incubation. Lovett *et al*. (1988) observed penetration of endothelial monolayers of 7.20% and 1.20% for *T. pallidum*, and *T. phagedenis,* respectively, over an incubation period of 6 hours. When *T. denticola* and *T. phagedenis* were added to HUVEC monolayers, their respective percent penetrations over 12 hours were 8% and close to 0% (Peters *et al*., 1999). Similarly, when incubated for 10 hours, the traversal percentages of *T. denticola, T. pallidum*, and *T. phagedenis* across a tissue layer of immortalized human keratinocytes were 8%, 14%, and close to 0% (Lux *et al*., 2001).

In the current assay, dissemination of *T. phagedenis* across the HUVEC monolayer seeded on top of the growth factor-reduced Matrigel was assessed after 48 hours of incubation at 25°C under anaerobic conditions. The media from the bottom compartments of the dissemination transwell chambers were sampled for dark-field microscopy counts and the rest were used for total genomic DNA extraction and real-time quantitative PCR analysis. Analogous methodology was applied to the wells that did not have permeable inserts and hence no artificial endothelial barriers. The same number of bacteria was added to these wells as to their corresponding transwell inserts and the no-insert wells thus served to provide the total number of bacteria for each strain at the end of the incubation period. The traversal ratio for each strain was thus established by dividing the averaged number of treponemes that crossed the artificial endothelial barrier by the averaged total number of bacteria present at the end of the assay. Comparison of traversal ratios as opposed to comparison of exact number of treponemes that penetrated

the barrier controlled for any potential variations in the replication or growth rates of the different strains. Transformation of *T. phagedenis* with the *pallilysin*/pKMR plasmid required electroporation. Such manipulations could have selected for a sub-population of *T. phagedenis* with different qualities, such as different resilience, motility, or growth rate, than those of wildtype *T. phagedenis.* Since the duration of the Lovett *et al.* experiment from 1988 was much shorter than the doubling time of the spirochetes under study, any differences between the growth rates of the different microorganisms used would have been insignificant. However, the current assay was performed over 48 hours, which is well beyond the typical doubling time of treponemes (about 30 hours). Therefore, a difference in the growth rate of the different *T. phagedenis* strains could influence the traversal count. For these reasons, for each transwell containing an artificial endothelial barrier, another one without one was used in the assay. Thus, there were two transwells with wildtype *T. phagedenis* and two regular wells with the same number of wildtype *T. phagedenis* added, and three transwells plus three regular wells with either of pKMR/*T.phagedenis* and *pallilysin*/pKMR/*T. phagedenis.*

Dark-field Microscopy

First I present the results from the dark-field microscopy counts. To determine the dissemination capacity by dark-field microscopy analysis, three independent counts, 5 fields of view each, were performed on each transwell well. Dissemination capacity was represented by the average percent traversal values for each sample (Figure 14).

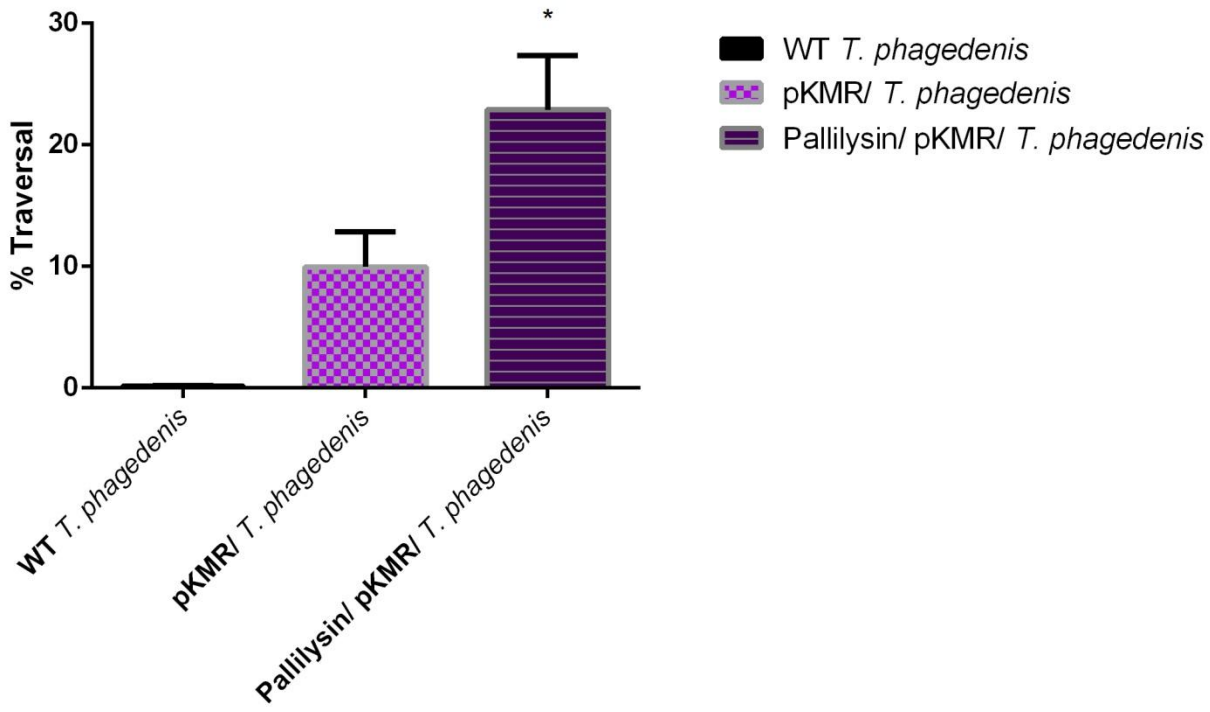


Figure 14. Heterologous expression of pallilysin conferred upon *T. phagedenis* the capacity to disseminate across an artificial endothelial barrier. Dissemination across a confluent HUVEC monolayer seeded on Matrigel-coated permeable inserts was assessed over 48 hours at 25°C under anaerobic conditions. Quantification was performed by dark-field microscopy: 3 counts/well with 5 fields of view each with each well treated as an independent sample.. Data were analyzed using one-way-ANOVA and Tukey's multiple comparison test to compare the average percent traversal values of wildtype *T. phagedenis* (n=2), pKMR/*T. phagedenis* (n=3), and pallilysin/pKMR/*T. phagedenis* (n=3), where n refers to the number of wells for each particular treatment; p=0.0177 pallilysin/pKMR/*T. phagedenis* vs. WT *T. phagedenis*; p=0.0892 pallilysin/pKMR/*T. phagedenis* vs. pKMR/*T. phagedenis*; p=0.2461 pKMR/*T. phagedenis* vs. WT *T. phagedenis;* graph shows mean + S.E.M.

Time restrictions due to the long period of optimization and contamination issues allowed only one experiment to be carried out. Therefore, the statistical analysis was performed on a single plate with each transwell of each strain treated as an independent sample as a proof of principle. Thus, based on the results from the one-way ANOVA analysis there was a statistically significant difference between the dissemination capacities of *T. phagedenis* heterologously expressing pallilysin and wildtype *T.*

phagedenis. Surprisingly, there was no statistically significant difference between the percent traversal values of pKMR/*T.phagedenis* and *pallilysin*/pKMR/*T. phagedenis* (p=0.0892), which raised the suspicion that the presence of the pKMR plasmid could be having an effect on the dissemination capacity of *T. phagedenis.* Alternatively, a potential selection effect could have occurred as a consequence of transforming *T. phagedenis* during electroporation; as mentioned earlier, electroporation could have selected for a sub-population of treponemes with increased fitness and motility. However, there was also no statistically significant difference between the percent traversal values of pKMR/ *T. phagedenis* and wildtype *T. phagedenis.* These results suggested that even assuming a selection event had indeed taken place, it was the presence of pallilysin on the surface of *T. phagedenis* that rendered significant the difference in dissemination between *pallilysin*/pKMR/*T. phagedenis* and wildtype *T. phagedenis.* However, the assay has to be repeated to satisfy the statistical requirement for independent samples.

Due to the small sample size, even if heterologous expression of pallilysin in *T. phagedenis* indeed conferred dissemination capacity upon *T. phagedenis,* that treatment might not have resulted in a statistically significant difference in the current assay. To address this possibility, power analysis was performed. Power is based on the sample size, variation in the experiment and the size of the difference between treatments tested. It represents the fraction of experiments expected to yield a statistically significant p-value given the afore-mentioned assay parameters. High power translates as a high chance that the experiment will find a statistically significant result if the treatment really has an effect. Statistical power analysis performed with StatMate software based on the darkfield microscopic counts indicated that the current assay had only a 50% power to

detect a significant difference (alpha = 0.05) between the current means of pKMR/*T.phagedenis* and *pallilysin*/pKMR/*T. phagedenis.* Commonly, experiments are designed to yield 80% power. Thus, in order to lower the chance of missing a statistical difference between pKMR/*T.phagedenis* and *pallilysin*/pKMR/*T. phagedenis*, based on the current difference in means (12.9 units), the experimental variation and power of 80%, the sample size should comprise five independent samples per treatment. A sample size of 5 in each group will thus have 80% power to detect a difference between means of 12.68 units with a significance level (alpha= 0.05).

The average percent traversal of wildtype *T. phagedenis* biotype Kazan (0.1%) was in agreement with the results from previous studies done with *T. phagedenis* biotype Reiter (Comstock and Thomas, 1989; Lovett *et al*., 1988; Lux *et al*., 2001; Peters *et al*., 1999; Riviere *et al*., 1989). However, when compared to the percent traversal of virulent *T. pallidum*, the traversal rate of *T. phagedenis* heterologously expressing pallilysin (roughly 23%) is much higher (Comstock and Thomas, 1989; Lux *et al*., 2001; Riviere *et al*., 1989). According to Haake and Lovett (1994), the rate at which spirochetes penetrated an endothelial monolayer was linear for the first 4 hours. In contrast, Lux *et al*. (2001) reported exponential penetration kinetics for *T. pallidum* and *T. denticola* over a period of 10 hours. Therefore, it is possible that the duration of the current assay was long enough (48 hours) for such an exponential increase in traversal rates to have occurred. To our knowledge, the longest incubation of *T. pallidum* with a cell monolayer was 10 hours with a resulting percent traversal of 14% (Lux *et al*., 2001). In addition, the researchers in this study also found that tissue penetration rates were also dependent on the tissue

resistance (TEER). Increased tissue resistance delayed the onset of detectable penetration; however, the invasion kinetics remained exponential. Therefore, certain events during the penetration must be facilitating and increasing the likelihood for other treponemes to cross the tissue layer. Lux *et al*. (2001) hypothesized that release of chemo-attractants as a result of tissue destruction could chemotactically target weakened spots thus facilitating further tissue penetration. However, in their study no significant loss of tissue resistance was observed to indicate massive tissue destruction.

In the process of optimizing the dissemination assay for the current study I also noticed that there was an exponential relationship between the HUVEC TEER and the traversal rates of the treponemes (Figure 15).

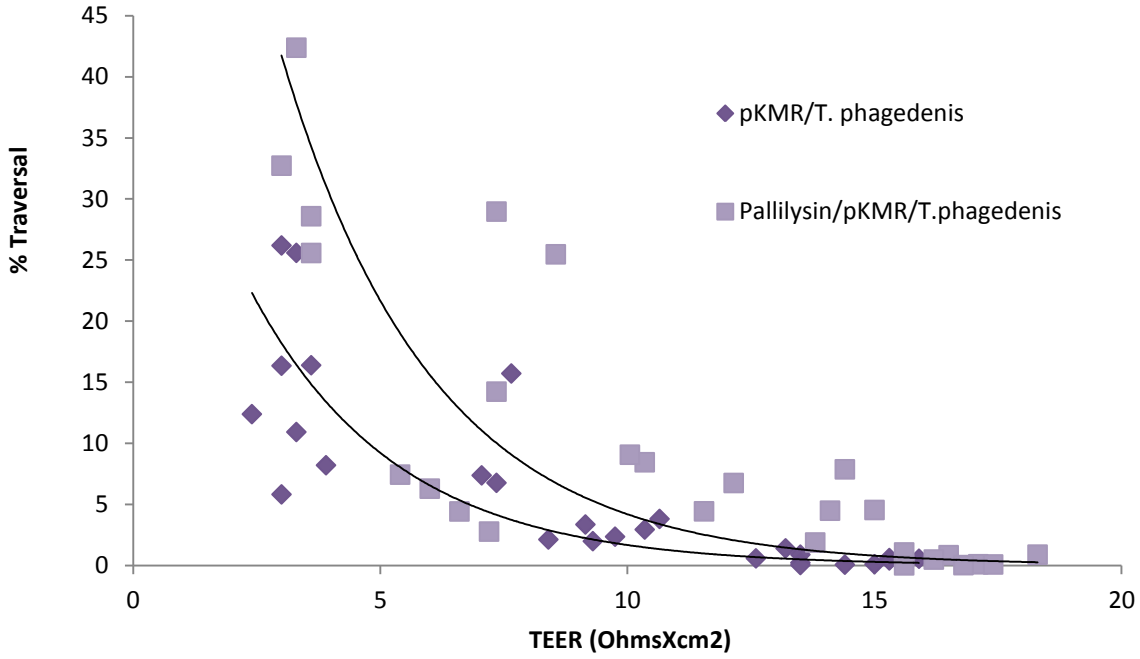


Figure 15. Correlation of penetration rate and transenddothelial electrical resistance (TEER). Traversal rates were determined after 30 to 48 hours of coincubation of the HUVEC monolayers with pKMR/ *T. phagedenis* and pallilysin/pKMR/ *T. phagedenis*. TEER presented was measured at the end of the experiment.

As seen in Figure 15, higher final TEER corresponded to lower percent traversal. The samples that exhibited final TEER values between 15 and 20 OhmXcm^2 showed a very low corresponding percent traversal. These samples were part of the initial stages of assay optimization and were incubated for 30 hours at 25°C. However, as discussed earlier, since pallilysin's activity is lowered at 25°C, the assay was extended to 48 hours.

It is important to note that the TEER of all samples had decreased over the 48-hour incubation period, including the one without treponemes (Figure 16). This was indicative of decrease in monolayer integrity due still in part to the culture conditions. However, the TEER remained over 10 OhmsXcm^2 for both HUVECs alone and HUVECs in the presence of wildtype *T. phagedenis*, indicative of a confluent monolayer (Haake and Lovett, 1994a). In contrast, the TEER values for the other two strains had dropped further by roughly 3 units. Since wildtype *T. phagedenis* in the current study (Figure 15) exhibited nearly no traversal at all in comparison with the other two strains, the HUVEC monolayer must have remained confluent enough to prevent non-specific traversal of this non-invasive treponeme. This in turn suggested that the traversal of the other two strains was not due to a non-specific event as the result of breach of monolayer integrity, but rather was treponeme- and strain-specific. Consistent with the higher traversal rate of pKMR/*T.phagedenis* and *pallilysin*/pKMR/*T.phagedenis* a drop in the respective TEER values of the monolayers exposed to these strains was observed Figures 15 and 16). An inverse correlation between traversal rates and change in TEER has been reported previously supporting the current results (Grab *et al*., 2009, 2013).

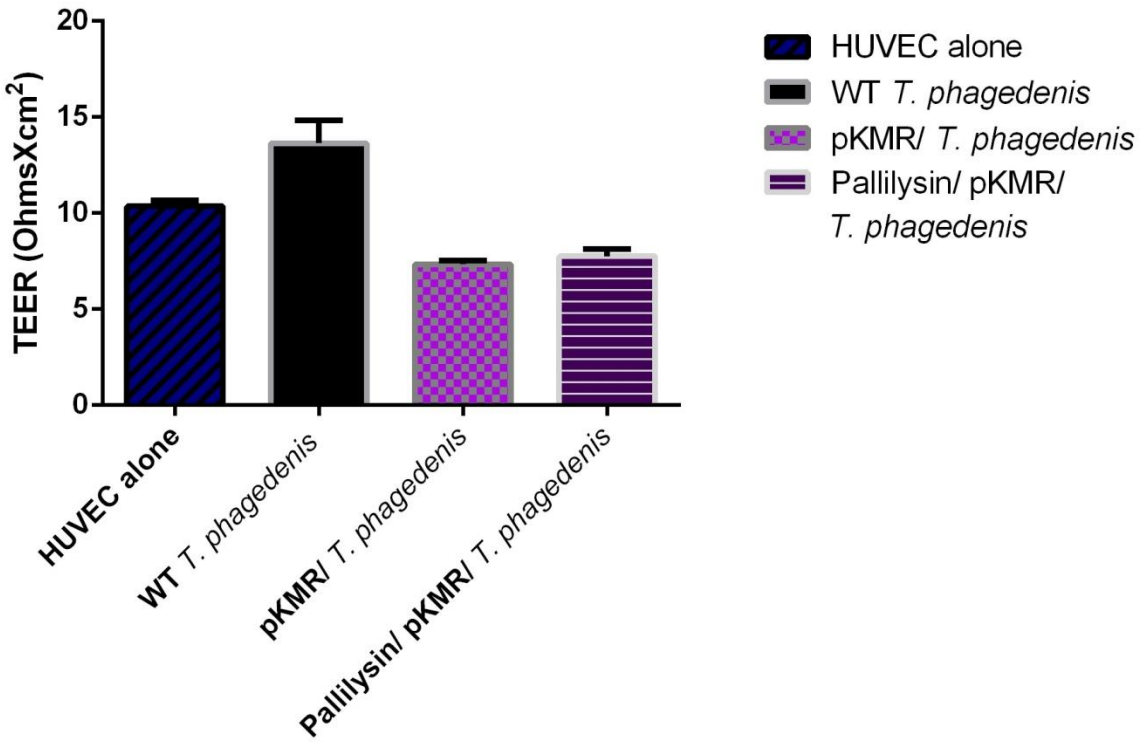


Figure 16. *T. phagedenis h*eterologously expressing pallilysin and pKMR/*T. phagedenis* adversely affect HUVEC monolayer integrity.
TEER was measured after incubation for 48 hours at 25ºC under anaerobic conditions. Bars represent mean transendothelial electrical resistance (TEER (ohmsXcm^2)) + S.E.M. of wildtype *T. phagedenis* (n=2), pKMR/*T. phagedenis* (n=3), and pallilysin/pKMR/*T. phagedenis* (n=3), where n refers to the number of wells for each particular treatment; Two wells were also seeded with HUVECs only. p=0.0296 HUVEC alone vs. WT *T. phagedenis;* p=0.0296 HUVEC alone vs. pKMR/*T. phagedenis*; p=0.0534 HUVEC alone vs. Pallilysin/pKMR/*T. phagedenis*; p=0.0007 WT *T. phagedenis vs.* pKMR/*T. phagedenis;* p=0.0010 WT *T. phagedenis vs.* Pallilysin/pKMR/*T. phagedenis*; p=0.9329 pKMR/*T. phagedenis vs.* Pallilysin/pKMR/*T. phagedenis*

Statistically significant differences were observed between the TEER values wildtype *T. phagedenis* and either pKMR/*T. phagedenis* or *pallilysin*/pKMR/*T. phagedenis* (Figure 16). However, wildtype *T. phagedenis* was not expected to have an effect on the TEER of the HUVEC monolayer, yet there appeared to be a statistically significant difference between the two means. Perhaps, even more surprisingly, there is

an increase of TEER when wildtype *T. phagedenis* were incubated with the HUVECs. This could possibly be due to clumping of the bacteria or contamination which could have artificially increased the TEER. Therefore, only relying on the TEER difference between wildtype and the other two strains might be misleading. When compared to the TEER value of HUVEC alone, there appeared to be a significant difference with pKMR/*T. phagedenis,* but not with *pallilysin*/pKMR/*T. phagedenis.* No statistically significant difference was observed between pKMR/*T. phagedenis* and *pallilysin*/pKMR/*T. phagedenis,* either. The power of the current design was, however, only 10%; hence, there was very little power to detect a statistically significant difference (α=0.05) between pKMR/*T. phagedenis* and *pallilysin*/pKMR/*T. phagedenis.*

Experimental sample size analysis based on the TEER values of pKMR/*T. phagedenis* and pallilysin/pKMR/*T. phagedenis* revealed that a sample size of 25 in each group will have 80% power to detect a statistically significance difference between the two means (alpha= 0.05).

.

Real-time quantitative PCR

Real-time quantitative PCR (QPCR) is a major analytical method in molecular biology that allows kinetic detection of accumulating PCR products over the cycling period. In addition, the fluorescence based detection provides greater sensitivity for amplicon detection as compared to conventional gel-based detection (Yun *et al*., 2006). Major applications of the QPCR methodology include mRNA expression studies (Li and Wang, 2000; Rose'Meyer *et al*., 2003; Tchaptchet *et al*., 2012) and SNP detection (Centurion-Lara *et al*., 2013), as well as cancer marker quantification (Mocellin *et al*.,

2003), genotyping (Ponchel *et al*., 2003), and pathogen load diagnosis (Mengelle *et al*., 2003; Polanco *et al*., 2002). As an extension of the latter, QPCR was used in the current study to quantify the dissemination capacity of *T. phagedenis* across the artificial endothelial barrier. Total genomic DNA was extracted from the treponemes found in the bottom compartment of the transwells, as well as from the treponemes in the regular wells that represented the total number of bacteria at the end of the incubation period. A region from the *T. phagedenis* flagellar hook polypeptide gene (*flgE*) was amplified in real time using SYBR Green I detection. The resulting *flgE* copy number from each transwell directly corresponded to the number of bacteria that had crossed the artificial endothelial barrier. However, to determine the gene copy number and thus the number of treponemes using QPCR a standard curve of known quantities of the PCR target sequence was generated. In QPCR, the cycle at which the amount of detectable PCR product (fluorescence based) reaches a preset threshold level (Ct) is assessed (Yun *et al*., 2006). A standard curve correlates the Ct readings to the known number of gene copy number present at the start of the QPCR analysis. This in turn facilitates the extrapolation of the unknown gene copy numbers based on their respective Ct readings from the standard curve. For this reason, a standard curve of initial *flgE* copy numbers from 10-fold serial dilutions of known *T. phagedenis* numbers ($1x10^8$ to $1x10^2$) was created (Figure 17). For example, a dilution containing $1x10^8$ treponemes corresponded to $1x10^8$ initial *flgE* copies on the standard curve.

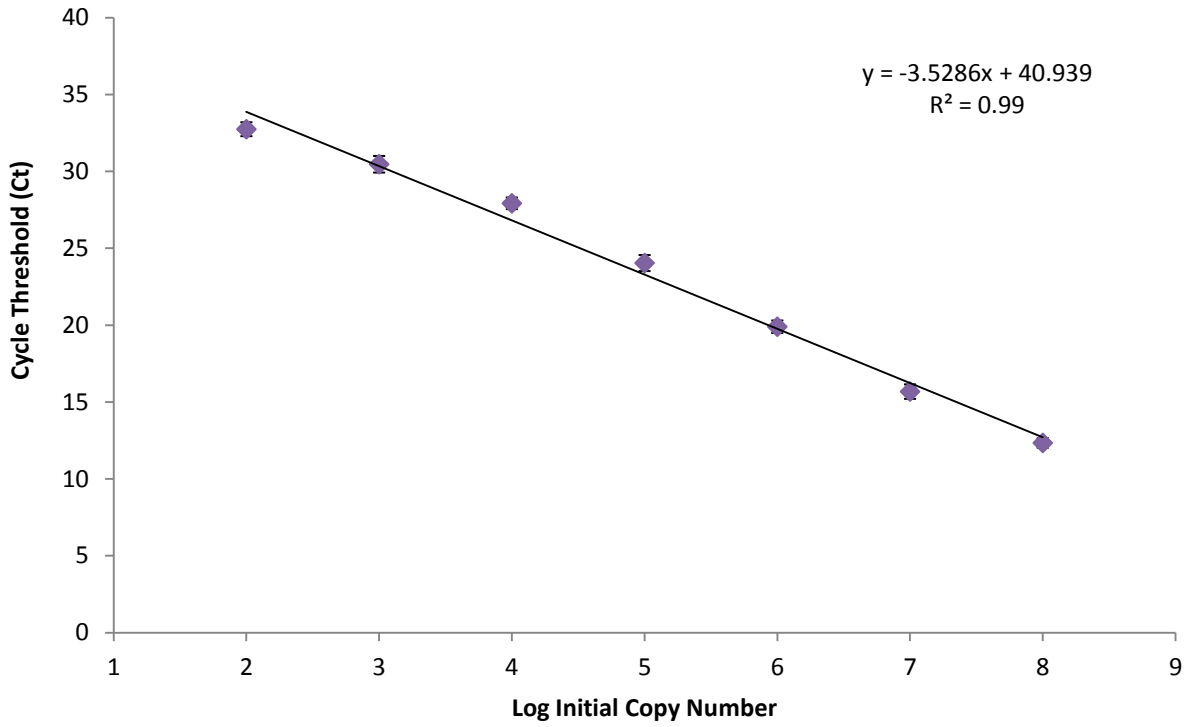


Figure 17. Standard curve generated to calculate initial *flgE* copy number in the quantification of treponemal dissemination. The threshold cycle value (Ct), indicative of the quantity of the *flgE* gene at which the fluorescence exceeds a pre-set threshold, was determined. The Ct values obtained through SYBR Green I detection were plotted against the log of the initial *flgE* copy number from triplicate 10-fold *T. phagedenis* serial dilutions ($1x10^8$ to $1x10^2$ treponemes). PCR reactions were performed in quadruplicate for 40 cycles (95°C for 15 seconds, 65°C for 20 seconds, and 72°C for 40 seconds) following initial Maxima Hot Start Taq DNA Polymerase activation at 95°C for 10 minutes. Melting curve analysis was carried in the end.

The equation of the line was then used to quantify dissemination. The samples from the dissemination assay plate were subjected to genomic DNA extraction and a region from the *flgE* gene was amplified in quadruplicate from each of wildtype *T. phagedenis* (n=2), pKMR/ *T. phagedenis* (n=3), and *pallilysin*/pKMR/*T. phagedenis* (n=3) and their corresponding total bacteria counterparts. A positive control from the standard curve was run together with the rest of the samples in triplicate to control for plate and reagent variability and ensure transferability of test and standard curve Ct's.

Variation of 1 Ct or less between plates was deemed acceptable. Negative controls were run in triplicate on the plate and included a no-primer, a no-template, and a no-DNA polymerase control. Melting curve analysis was performed at the end of the QPCR run to verify the specificity of each reaction. As seen from Figure 19, the melting curves for the individual reactions were tightly clustered together with narrow peaks and average melting temperature (Tm) of 86.3°C, median Tm of 86.3°C, and average standard deviation of only 0.29°C, indicative of highly specific amplification.

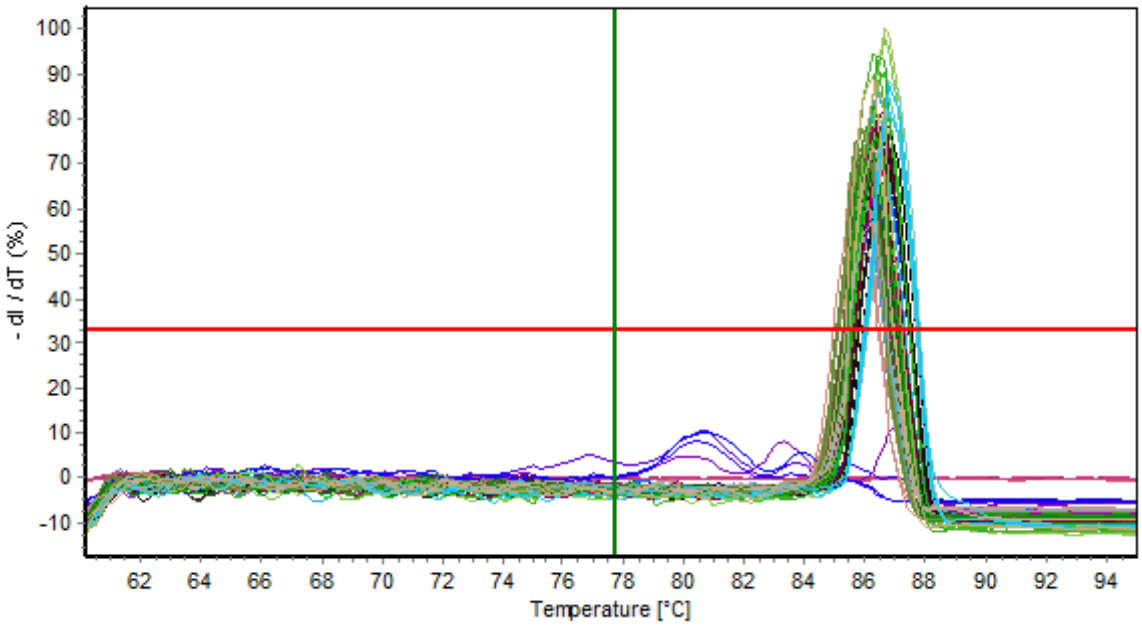


Figure 18. Melting curve analysis of QPCR reactions from the dissemination assay. Threshold was set at 33% and the average melting temperature including that of the positive control was 86.3°C. No melting curves were generated from the negative controls.

Analysis of the Ct's from the dissemination assay samples revealed a high degree of amplification precision within each of the quadruplicate samples with an average standard deviation of only 0.27 Ct's. While there was significant variation among dark-field microscopy counts from the same sample, the variability in the process of

quantification of dissemination was drastically minimized by using the QPCR methodology. This was a definite step towards lowering the technical variation introduced in the process of bacterial enumeration. However, even with more precise methods for quantifying traversal rates, the fact that *T. phagedenis* is prone to clumping remains. Coupled with the fragile outer membrane of the treponemes, preventing thorough mixing, there was a higher chance of pipetting errors. That in turn could have given rise to technical variability during assay set up. Indeed, analysis of the initial *flgE* copy numbers from the wells without endothelial barriers revealed discrepancies between samples that were supposed to have had the same number of bacteria added. Therefore, the variation in dissemination observed between transwells of the same strain of *T. phagedenis* could have in part been due to technical reasons. As seen in Figure 19, the dissemination trend was conserved between the two quantification methods. No statistically significant difference was observed between wildtype *T. phagedenis* and *T. phagedenis* expressing pallilysin, nor between pallilysin/pKMR/*T. phagedenis* and pKMR/*T. phagedenis*. However, the technical error in setting up the dissemination assay likely increased the perceived variability and contributed to the lack of significant difference in dissemination. More replicates are necessary to differentiate technical noise from real biological effects. However, the fact that the same trend was observed after both methods for assessing traversal supported a potential role of pallilysin in conferring dissemination capacity to *T. phagedenis.* However, the results are inconclusive and more experimentation with a different sample size is necessary.

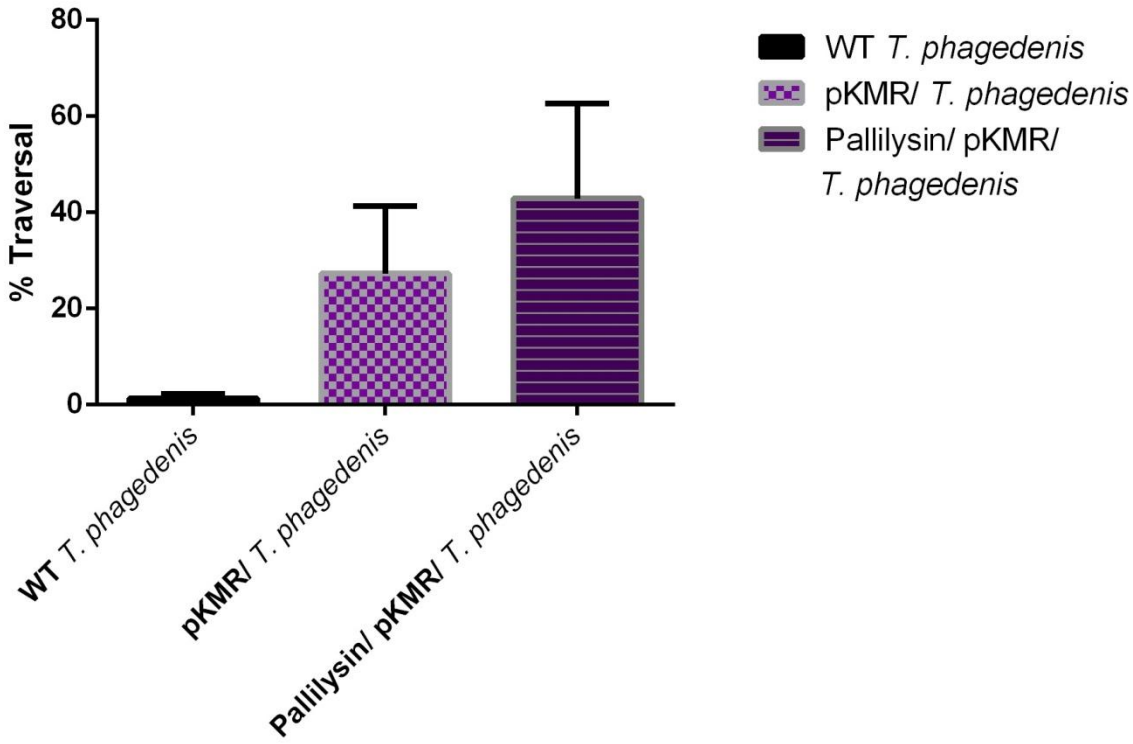


Figure 19. Percent traversal of *T. phagedenis* across the artificial endothelial barrier quantified with real-time quantitative PCR. Dissemination of *T. phagedenis* across a confluent HUVEC monolayer seeded on Matrigel coated permeable inserts was assessed over 48 hours at 25°C under anaerobic conditions. Quantification of traversal ratios was based on the amplification of a region in the *flgE* gene from *T. phagedenis* performed on real-time quantitative PCR. The samples were amplified in quadruplicate and detected with SYBR Green I. Standard curve was used to extrapolate the initial *flgE* copy numbers in each reaction based on the respective fluorescent readings. The numbers of bacteria directly proportional to the initial copy numbers were used to calculate traversal ratios. Data was analyzed using one-way-ANOVA and Tukey's multiple comparison test to compare the average percent traversal values of wildtype *T. phagedenis* (n=2), pKMR/*T. phagedenis* (n=3), and *pallilysin*/pKMR/*T. phagedenis* (n=3), where n refers to the number of wells for each particular treatment; p=0.2874 pallilysin/pKMR/*T. phagedenis* vs. WT *T. phagedenis*; p=0.5695 pKMR/*T. phagedenis* vs. WT *T. phagedenis*; p=0.7625 pallilysin/pKMR/*T. phagedenis* vs. pKMR/*T. phagedenis;* graph shows mean + S.E.M.

The power analysis revealed that the current statistical power of the experiment quantified by qPCR was under 10%. In order to observe a statistically significant

difference between pKMR/*T.phagedenis* and *pallilysin*/pKMR/*T. phagedenis* based on the current difference in means (12.7 units), with the current experimental variation and a power of 80%, the sample size should comprise two hundred independent samples per treatment. The reason for such a large sample size are mainly two factors: there was high experimental variability (average standard deviation of 45 units) and the small difference between the two means requires a larger sample size to increase the probability of observing a significant p-value. However, provided that the technical variability is reduced in setting up the assay, the standard deviation would likely be reduced and so would the estimated sample size to a number closer to the sample size based on the dark-field microscopic counts (5).

The exact mechanism of endothelial monolayer invasion by spirochetes, and in particular by *T. pallidum*, remains unknown. HUVEC activation by virulent *T. pallidum* increased expression of ICAM-1 and procoagulant activity on the surface of the endothelial cells with extensive networks of fibrin strands (Riley *et al*., 1992). Whether activation of endothelial cells by *T. pallidum* promotes tissue invasion by the pathogen remains unclear. It should be noted, however, that pallilysin is capable of degrading fibrin clots (Houston *et al*., 2012).

In studies of *T. pallidum* invasion, an abundance of organisms were found in the intercellular junctions between endothelial cells, unlike the non-pathogenic *T. phagedenis* biotype Reiter (Haake and Lovett, 1994a; Lovett *et al*., 1988). Intercellular *T. pallidum* were observed in close proximity to tight junctions, but the TEER readings remained unaltered throughout the duration of the experiment indicating that the intercellular junction integrity of the endothelial monolayer had been maintained (Lovett *et al*., 1988).

In contrast, the current dissemination assay study showed clear correlation between increased traversal and decreased transendothelial resistance. The exact mechanism of traversal thus remains controversial. Interestingly, previous research also found intracellular treponemes (Haake and Lovett, 1994b). When either *T. pallidum* or the non-pathogen *T. phagedenis* biotype Reiter were exposed to endothelial monolayers some of them appeared to have been phagocytosed by the endothelial cells. A transcellular route of traversal through the monolayer, and in particular one that does not damage the cells, could explain the lack of decrease in TEER readings observed in the study by Lovett *et al*. (1988). Alternatively, if the passage was paracellular, *T. pallidum* must possess a specific mechanism for opening tight junctions without damaging them. The mechanism could be similar to regulated opening and closure of tight junctions during leukocyte extravasation. In concert with the latter notion, thrombin, among its other roles in the coagulation cascade, exhibits pro-inflammatory properties (Vergnolle *et al*., 2002). Thrombin induces cytokine release, leukocyte rolling and adherence to the vascular endothelium, and also increases vascular permeability. In addition, pallilysin can be proteolytically activated by thrombin. Therefore, it is possible that pallilysin is aided by thrombin in both its activation and interaction with the tight junctions either directly or indirectly to permit transendothelial traversal of *T. pallidum.*

4.4 Conclusions

In order to test whether heterologous expression of pallilysin conferred upon *T. phagedenis* a dissemination capacity, a transwell dissemination assay was designed. Percent traversal across the artificial endothelial barrier was assessed for each of wildtype

T. phagedenis, pKMR/*T.phagedenis*, and *pallilysin*/pKMR/*T. phagedenis* using dark-field microscopy counts and qPCR.

Extensive optimization of the assay design was necessary. To remain viable for the duration of the assay *T. phagedenis* had to be cultured under anaerobic conditions, which proved to be less than optimal for the survival of the HUVECs. A special media mix was created and an exogenous buffering agent was added to maintain physiological pH under anaerobic conditions. In addition, the concentration of serum was also increased to help keep the HUVEC healthier. Furthermore, the incubation temperature had to be decreased to 25°C to maintain HUVEC monolayer integrity under anaerobic conditions. The temperature drop, however, also lowered the activity of pallilysin. The assay thus had to be performed over 48 hours.

There was a clear correlation between the percent traversal of the treponemes and the decrease in TEER readings at the end of the experiment, indicative of loss of monolayer integrity. An overall decrease in TEER readings was observed including those for the HUVEC monolayer alone. However, either of pKMR/ *T. phagedenis* or *pallilysin*/pKMR/*T. phagedenis* had a further negative effect on the monolayer integrity. Dark-field microscopy counts revealed that *T. phagedenis* heterologously expressing pallilysin had the highest traversal rate, followed by pKMR/ *T. phagedenis.* Wildtype *T. phagedenis* had traversal rates close to zero. Statistical analysis performed on the dissemination counts from a single plate as proof of principle revealed a significant difference between the dissemination capacities of *T. phagedenis* heterologously expressing pallilysin and wildtype *T. phagedenis.* However, there was no significant difference between the means of pallilysin/pKMR/*T. phagedenis* and pKMR/*T.*

phagedenis. Such a statistically significant difference is more important than that between wildtype *T. phagedenis* and pallilysin/pKMR/*T. phagedenis,* because of possible selection or plasmid effects. Further analyses are necessary with an increased sample size.

To circumvent problems with the variation in the quantification methodology due to the nature of the treponemes, qPCR analysis was also performed on the dissemination assay samples. A standard curve was created from known dilutions of *T. phagedenis* to calculate the number of bacteria in each dissemination assay sample from the respective Ct readings produced. The QPCR methodology proved to be very precise within the quadruplicate replica samples in contrast to the variation observed with the replica samples from the dark-field microscopy counts. However, technical variation from mixing and pipetting errors early in the assay set up due to clumping of the treponemes likely contributed to the variability observed between replica transwells of the same strain. Even with such high variability, the dissemination trend was conserved; however, no significant difference was observed among any of the treatments. The power of the QPCR assay was extremely low (more than 5 times lower than that of dark-field microscopic counts) and as such there was a very limited chance of observing a statistically significant difference between the means of pallilysin/pKMR/*T. phagedenis* and pKMR/*T. phagedenis.* However, based on the higher level of technical precision in quantifying dissemination, QPCR proved to be the better method for future studies. A significant sample size increase (200) would be necessary to boost the power of the assay to 80%, but with less technical variation the sample size could drop down to around that for the dark-field microscopic counts (5). Table 5 summarizes the current sample size and

power of the assays and also lists the desired sample sizes for each assay based on power analyses. Only calculations for the difference between the means of pallilysin/pKMR/*T. phagedenis* and pKMR/*T. phagedenis* were included, since such a difference would indicate a role of pallilysin in *T. pallidum* crossing of the endothelial barrier.

Table 5. Power and sample size calculations for dissemination of *T. phagedenis* heterologously expressing pallilysin

Assay Type	Current Power	Current sample size (wells)		Minimal Sample Size; Power=80%	
		pKMR/ *T. phagedenis*	Pallilysin/pKMR/ *T. phagedenis*	pKMR/ *T. phagedenis*	Pallilysin/pKMR/ *T. phagedenis*
Dark-field Microscopy	50%	3	3	5	5
QPCR	<10%	3	3	200	200
TEER	10%	3	3	25	25

In summary, my thesis research results from the dissemination and TEER assays do not provide conclusive evidence for a role of pallilysin in disseminating across an endothelial barrier. However, the dissemination trend obtained from the dark-field microscopic counts and QPCR was conserved. With an increase in sample size and thus power of the assays, detection of statistically significant difference between pallilysin/pKMR/*T. phagedenis* and pKMR/*T. phagedenis* will be more likely. Such results would be in concert with the overwhelming mass of evidence, previously obtained in our laboratory, which implicated pallilysin with a role in the dissemination of *T. pallidum* (Cameron, 2003; Cameron et al., 2005, 2008b; Houston et al., 2011, 2012).

Chapter 5: Significance and Future Directions

5.1 Significance

In the past, research in the molecular basis of *T. pallidum* pathogenesis and dissemination has been severely hindered due to a number of factors. Firstly, the inability to continuously culture the pathogen *in vitro* has limited the number of studies that could be performed to elucidate *T. pallidum* pathogenesis (Fieldsteel *et al*., 1981, 1982; Norris, 1982). Secondly, the pathogen is genetically intractable, which precluded researchers from directly investigating the functions of its gene products through traditional genetic manipulations. Thirdly, its outer membrane is easily disrupted by conventional laboratory manipulations, such as centrifugation and resuspension (Bailey et al., 1985; Radolf et al., 1989). Finally, the very low outer membrane protein content and the limited sensitivity of conventional technologies limited the identification of *T. pallidum* OMPs, which directly contact the host and are most likely to mediate dissemination (Walker et al., 1989). Our laboratory identified pallilysin as a surface-exposed adhesin and protease, capable of binding to and degrading laminin, a major component of basement membranes (Cameron, 2003; Cameron et al., 2005; Houston et al., 2011, 2012). These and the rest of its spectrum of activities strongly suggested a role in the dissemination of *T. pallidum.*

The current bioinformatics analyses on the pallilysin sequence confirmed that the HEXXH motif, crucial for proteolysis, was unique and only present in the homologues from *T. pallidum* and *T. paraluiscuniculi.* Both species cause infections with disseminated patterns in humans and rabbits, respectively. Apart from the HEXXH motif, the analyses showed that the pallilysin sequence itself, and in particular the mid to C-terminal region, was relatively well conserved among homologues from the *Treponema*,

Spirochaeta, and to a lesser extent from the *Borrelia* genera. Therefore, the protein itself was not an independent invention of *T. pallidum*, but shares evolutionary ties to homologues in related spirochetes. The amino acid conservation analysis pointed to an alternate unknown function of the *T. pallidum* protease, one that would be conserved among the studied homologues. Interestingly, however, the branch length on the phylogenetic tree and the amino acid frequency analysis of the *T. pallidum* and the *T. paraluiscuniculi* homologues indicated that these two sequences were well separated from the rest of the homologues. A striking example was the abundance of histidine (H) residues in the pallilysin sequence, which may have facilitated the *de novo* formation of the unique HEXXH motif contributing to the pathogenicity of *T. pallidum.* In conclusion, the comparison of pallilysin in the context of its homologues further solidified a role in the dissemination of *T. pallidum*.

T. pallidum is one of the most invasive human pathogens known, able to rapidly spread through the bloodstream and establish a disseminated infection. Hematogenous spread, however, requires that the pathogen cross the endothelium lining the blood vessels. Therefore, *T. pallidum* must have a specific mechanism to overcome the tight junctions that seal the endothelial cells into a virtually impenetrable barrier. Indeed, *T. pallidum* has been found in the interjunctional space between endothelial cells and shown to traverse an endothelial layer *in vitro* suggesting a paracellular route of dissemination (Thomas *et al*., 1989). However, since this and similar studies from the 1980's no particular overarching molecular mechanism of dissemination has been described so far.

To circumvent the limitations surrounding *T. pallidum* research, and based on the accumulated evidence for a role of pallilysin in dissemination, the current study aimed to

investigate whether heterologous expression of pallilysin conferred upon *T. phagedenis* the capacity to disseminate across an artificial endothelial barrier. The assay utilized the non-pathogenic model treponeme to allow for the isolated study of pallilysin and its direct role in *T. pallidum* dissemination. A trend of higher percent traversal across the artificial endothelial barrier by *T. phagedenis* heterologously expressing pallilysin was observed in comparison to wildtype and pKMR/ *T. phagedenis.* The trend was conserved in both methods of assessment of dissemination capacity, namely dark-field microscopy counts and QPCR. Further experimentation is necessary to prove a significant difference of traversal between *T. phagedenis* heterologously expressing pallilysin and *T. phagedenis* carrying the empty shuttle vector. Confirmation of this trend and significance at the latter level would directly implicate, for the first time, a *T. pallidum* OMP in the dissemination of the pathogen specifically across the endothelium. That would in turn prompt investigations into the exact molecular mechanism of action by which pallilysin may facilitate the traversal of *T. pallidum* across the endothelial lining of blood vessels. Blocking of the functioning of this protein may help prevent the widespread dissemination of *T. pallidum.* Finally, based on its surface localization, expression during infection, and role in dissemination, pallilysin would present a good vaccine target candidate. The importance of such a vaccine stems from the fact that even though syphilis can be cured with penicillin, the disease remains a global public health concern with close to11 million new infections per year.

5.2 Future Directions

Further optimization should aim at maintaining monolayer integrity for the full duration of the assay under anaerobic incubation. Additional increase of the serum and

growth factor concentrations in the medium might stabilize the monolayer for longer. However, the endothelial cells have to be observed carefully for any signs of tube-formation due to the increase of growth factors.

In order to bring out the true biological variation and minimize technical variation, a slightly modified system for quantifying the traversal ratios should be implemented. So far the number of bacteria that crossed the barrier in the transwell was divided by the averaged total number of bacteria found in the corresponding wells without filters. The assumption that there would be little well to well variation within the same sample was proved erroneous by the QPCR analysis. Bacterial clumping was deemed the major reason for the observed variability. Ideally the ratios would have been calculated by dividing the number of bacteria found in the bottom compartment of the transwell by the sum of bacteria found in the top and bottom compartments of the same transwell. However, early in the optimization process this proved to be an unreliable method for the dark-field microscopy counts because the treponemes were attaching to, and clumping over, the Matrigel. Thus, inaccurate counts were obtained from the media in the top compartments. However, with the introduction of the QPCR that issue could be solved. A non-enzymatic recovery solution exists that should allow dislodging of the cells from the Matrigel. The Corning Cell Recovery Solution depolymerizes BD Matrigel Matrix gels without enzymatic digests and the cells are released without damage. Once the Matrigel is depolymerised, the treponemes should also be released. Total genomic DNA could then be extracted not only from the compartments below the filters, but also from the ones above. This could thus prove an effective technique for analysis with QPCR since the method relies on specific amplification of DNA. Then, traversal ratios

would be established from the same transwell, by dividing the number of treponemes found below the filter by the total number of bacteria found below and above the filter. Thus, variation due to using a separate regular well without an insert for the total number of bacteria should be significantly lowered.

To further strengthen the power of the assay, control transwells with permeable filters but without the endothelial monolayer or Matrigel should be included. In theory, all three strains should traverse the permeable filters at equal rates due to the lack of a significant barrier that would normally be provided by the Matrigel and the HUVEC monolayer. Traversal across the bare permeable filters would form a baseline for unimpeded traversal. Equal traversal across the bare filters would verify the difference in the observed penetration of the artificial endothelial barrier, i.e. almost no traversal for wildtype *T. phagedenis* compared to higher levels of penetration by the other two strains. Thus the possibility of wildtype *T. phagedenis* not traversing due to clumping, for example, and not due to the lack of pallilysin would be excluded.

Another way to verify the traversal rates among the strains would be to compare invasion in the presence or absence of EDTA (Haake and Lovett, 1994a). EDTA reversibly lowers the TEER of endothelial monolayers by opening intercellular junctions. Therefore, all three stains should be able to penetrate the endothelial monolayers in the presence of EDTA. In addition, EDTA chelates metal cations, including zinc, and would thus inhibit pallilysin, a zinc metalloprotease. This should effectively turn *pallilysin*/pKMR/*T.phagedenis* into pKMR/*T. phagedenis.*

A slight twist on the above method would involve the use of thrombin instead of EDTA. While thrombin would also disturb the monolayer integrity, hence it was not used

in the first round of optimization of the assay, it would also catalyze the activation of pallilysin. Therefore, with the right concentration and timing, such that the effect of pallilysin activation overpowers thrombin's effect on the monolayer, it might be possible to see a more pronounced difference in traversal between *pallilysin*/pKMR/*T.phagedenis* and pKMR/*T. phagedenis* much earlier than 48 hours. This would in turn increase the likelihood of survival for the HUVECs under anaerobic conditions.

Currently, the TEER was measured before and after the assay, which only provides the end points but no picture of the TEER dynamics throughout the whole length of the experiment. The study would benefit from measurements of the TEER in real time, which would enable a direct correlation between a specific spike, change or trend in the TEER and the corresponding traversal rate at that point. An apparatus exists that performs real time measurements of electric impedance called electric cell substrate impedance sensing (ECIS) (Grab *et al*., 2009).

In addition, to confirm that the increase of permeability, i.e. decrease of TEER readings, was not due to loss of cell viability, future investigations into the mechanism of *T. pallidum* invasion of endothelial monolayers could benefit from cell viability or cytotoxicity assays. Trypan blue exclusion or MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, are commonly used to measure cell viability (Sigma-Aldrich, 2011). The reduction of tetrazolium salts such as MTT to colored formazan compounds only occurs in metabolically active cells. Trypan blue, on the other hand, is excluded from live cells, while dead cells with permeated membranes take up the dye and are colored blue.

Potential toxic effects of recombinant pallilysin on the HUVEC monolayer as a means for transendothelial traversal could be evaluated. Analogous investigations could be performed with *T. phagedenis* heterologously expressing pallilysin. The ECIS can also be used for such purposes. Another example of a cytotoxicity assay is the lactate dehydrogenase (LDH) assay (Sigma-Aldrich, 2013). Lactate dehydrogenase is an oxidoreductase enzyme that catalyses the interconversion of pyruvate and lactate. Thus, LDH is released by cells into the bloodstream after tissue damage. Quantification of LDH can be used to evaluate the presence of damage and toxicity to the endothelial monolayer.

Testing the traversal rates of the three strains of *T. phagedenis* on different cell lines could provide further insight into the role of pallilysin in *T. pallidum* dissemination. Although HeLa cells are not a biologically relevant cell line, because they are a cancer cell line, they should perform better under anaerobic condition. In addition, they do not possess tight junctions (Cunniffe *et al*., 2012; Thomas *et al*., 1989); hence, it would be valuable to assess the dissemination rates of *T. phagedenis* across HeLa layers as well and compare to those across the primary endothelial monolayer. A contrasting scenario of a primary cell line with very strong tight junctions and thus high TEER readings would be another interesting comparison. Brain microvascular endothelial cells have TEER values of 1500–2000 $OhmsXcm^2$ and are also a biologically relevant cell line since *T. pallidum* is able to cross the blood-brain barrier (Lafond and Lukehart, 2006).

Finally, *T. pallidum* was shown to activate HUVECs and induce expression of the adhesion molecules ICAM-1 and E-selectin (Haake and Lovett, 1994a). Activation was also observed by the *T. pallidum* lipoprotein TpN47, but not by heat-killed *T. pallidum* or the non-pathogenic treponeme *T. phagedenis* (Lafond and Lukehart, 2006). Therefore,

endothelial cell activation is likely a pathogen-specific, active process mediated by specific *T. pallidum* molecules. Thus, HUVEC activation and corresponding changes in the expression levels of adhesion molecules by pallilysin and *T. phagedenis* heterologously expressing pallilysin could be investigated. Finally, a potential correlation between the endothelial cell activation and the percent traversal of pallilysin/pKMR/T. *phagedenis* could be tested as a possible mechanism of dissemination of *T. pallidum*.

Bibliography

Abt, B., Göker, M., Scheuner, C., Han, C., Lu, M., Misra, M., Lapidus, A., Nolan, M., Lucas, S., Hammon, N., et al. (2013). Genome sequence of the thermophilic fresh-water bacterium Spirochaeta caldaria type strain (H1(T)), reclassification of Spirochaeta caldaria, Spirochaeta stenostrepta, and Spirochaeta zuelzerae in the genus Treponema as Treponema caldaria comb. nov., Trep. Stand. Genomic Sci. *8*, 88–105.

Alderete, J.F., and Baseman, J.B. (1980). Surface characterization of virulent Treponema pallidum. Infect. Immun. *30*, 814–823.

Altschul, S.F., Madden, T.L., Schäffer, a a, Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.

Amerongen, G.P.V.N., Draijer, R., Vermeer, M. a., and van Hinsbergh, V.W.M. (1998). Transient and Prolonged Increase in Endothelial Permeability Induced by Histamine and Thrombin : Role of Protein Kinases, Calcium, and RhoA. Circ. Res. *83*, 1115–1123.

Anderson, J., and Van, C. (1995). Tight junctions of paracellular and the molecular permeability basis for regulation. Am. Physiol. Soc. *95*, 467–475.

Antal, G.M., Lukehart, S.A., and Meheus, A.Z. (2002). The endemic treponematoses. Microbes Infect. *4*, 83–94.

Arakawa, S., and Kuramitsu, H.K. (1994). Cloning and sequence analysis of a chymotrypsinlike protease from Treponema denticola. Infect. Immun. *62*, 3424–3433.

Bailey, M.J., Penn, C.W., and Cockayne, A. (1985). Evidence for the presence of lipopolysaccharide in Treponema phagedenis (biotype Reiterii) but not in Treponema pallidum. FEMS Microbiol. Lett. *27*, 117–121.

BC Centre for Disease Control (2013). Infectious Syphilis among gay , bisexual and other men who have sex with men in British Columbia 2003 to 2012.

BD Biosciences (2011). BD Matrigel matrix. 1–7.

Bertot, G.M., Restelli, M. a, Galanternik, L., Aranibar Urey, R.C., Valvano, M. a, and Grinstein, S. (2007). Nasal immunization with Burkholderia multivorans outer membrane proteins and the mucosal adjuvant adamantylamide dipeptide confers efficient protection against experimental lung infections with B. multivorans and B. cenocepacia. Infect. Immun. *75*, 2740–2752.

Blanco, D.R., Reimann, K., Skare, J., Champion, C.I., Foley, D., Exner, M.M., Hancock, R.E., Miller, J.N., and Lovett, M. a (1994). Isolation of the outer membranes from Treponema pallidum and Treponema vincentii. J. Bacteriol. *176*, 6088–6099.

Bodnar, a G., Cooper, J.M., Leonard, J. V, and Schapira, a H. (1995). Respiratory-deficient human fibroblasts exhibiting defective mitochondrial DNA replication. Biochem. J. *305 (Pt 3*, 817–822.

Brinkman, M.B., McGill, M. a, Pettersson, J., Rogers, A., Matejková, P., Smajs, D., Weinstock, G.M., Norris, S.J., and Palzkill, T. (2008). A novel Treponema pallidum antigen, TP0136, is an outer membrane protein that binds human fibronectin. Infect. Immun. *76*, 1848–1857.

Buchacz, K., Patel, P., Taylor, M., Kerndt, P.R., Byers, R.H., Holmberg, S.D., and Klausner, J.D. (2004). Syphilis increases HIV viral load and decreases CD4 cell counts in HIV-infected patients with new syphilis infections. AIDS *18*, 2075–2079.

Calvani, M., Rapisarda, A., Uranchimeg, B., Shoemaker, R.H., and Melillo, G. (2006). Hypoxic induction of an HIF-1alpha-dependent bFGF autocrine loop drives angiogenesis in human endothelial cells. Blood *107*, 2705–2712.

Cameron, C.E. (2003). Identification of a Treponema pallidum Laminin-Binding Protein. Infect. Immun. *71*, 2525–2533.

Cameron, C., and Lukehart, S. (2013a). Current status of syphilis vaccine development: Need, challenges, prospects. Vaccine *32*, 1602–1609.

Cameron, C.E., and Lukehart, S. a (2013b). Current status of syphilis vaccine development: Need, challenges, prospects. Vaccine *32*, 1602–1609.

Cameron, C.E., Brown, E.L., Kuroiwa, J.M.Y., Schnapp, L.M., and Brouwer, N.L. (2004). Treponema pallidum Fibronectin-Binding Proteins. J. Bacteriol. *186*, 7019–7022.

Cameron, C.E., Brouwer, N.L., Tisch, L.M., and Kuroiwa, J.M.Y. (2005). Defining the Interaction of the Treponema pallidum Adhesin Tp0751 with Laminin. Infect. Immun. *73*, 7485–7494.

Cameron, C.E., Kuroiwa, J.M.Y., Yamada, M., Francescutti, T., Chi, B., and Kuramitsu, H.K. (2008a). Heterologous expression of the Treponema pallidum laminin-binding adhesin Tp0751 in the culturable spirochete Treponema phagedenis. J. Bacteriol. *190*, 2565–2571.

Cameron, C.E., Kuroiwa, J.M.Y., Yamada, M., Francescutti, T., Chi, B., and Kuramitsu, H.K. (2008b). Heterologous expression of the Treponema pallidum laminin-binding adhesin Tp0751 in the culturable spirochete Treponema phagedenis. J. Bacteriol. *190*, 2565–2571.

Canale-Parola, E., and Breznak, J. (1975). Morphology and physiology of Spirochaeta aurantia strains isolated from aquatic habitats. Arch. Microbiol. *105*, 1–12.

Carreau, A., El Hafny-Rahbi, B., Matejuk, A., Grillon, C., and Kieda, C. (2011). Why is the partial oxygen pressure of human tissues a crucial parameter? Small molecules and hypoxia. J. Cell. Mol. Med. *15*, 1239–1253.

Centers for Disease Control and Prevention (2006). Sexually Transmitted Diseases Treatment Guidelines , 2006. MMWR Recomm Rep *55*, 1–94.

Centurion-Lara, a, Castro, C., Barrett, L., Cameron, C., Mostowfi, M., Van Voorhis, W.C., and Lukehart, S. a (1999). Treponema pallidum major sheath protein homologue Tpr K is a target of opsonic antibody and the protective immune response. J. Exp. Med. *189*, 647–656.

Centurion-Lara, A., Giacani, L., Godornes, C., Molini, B.J., Brinck Reid, T., and Lukehart, S. a (2013). Fine analysis of genetic diversity of the tpr gene family among treponemal species, subspecies and strains. PLoS Negl. Trop. Dis. *7*, e2222.

Chi, B., Chauhan, S., and Kuramitsu, H. (1999). Development of a system for expressing heterologous genes in the oral spirochete Treponema denticola and its use in expression of the Treponema pallidum flaA gene. Infect. Immun. *67*, 3653–3656.

Chung, K.Y., Kim, K.-S., Lee, M.G., Chang, N.S., and Lee, J.B. (2002). Treponema pallidum induces up-regulation of interstitial collagenase in human dermal fibroblasts. Acta Derm. Venereol. *82*, 174–178.

Cines, D.B., Pollak, E.S., Buck, C.A., Loscalzo, J., Zimmerman, G.A., Rodger, P., Pober, J.S., Wick, T.M., Konkle, B.A., Schwartz, B.S., et al. (2014). Endothelial Cells in Physiology and in the Pathophysiology of Vascular Disorders. Blood *91*, 3527–3561.

Claude, P. (1978). Morphological factors influencing transepithelial permeability: a model for the resistance of the zonula occludens. J. Membr. Biol. *39*, 219–232.

Collart, P., Franceschini, P., and Durel, P. (1971). Experimental rabbit syphilis. Brit. J. Vener. Dis. *47*, 389–400.

Communicable Disease Prevention BC CDC (2013). British Columbia Annual Summary of Reportable Diseases - 2012.

Comstock, L.E., and Thomas, D.D. (1989). Penetration of endothelial cell monolayers by Borrelia burgdorferi. Infect. Immun. *57*, 1626–1628.

Cox, D.L., Chang, P.O., Mcdowall, A.W., and Radolf, J.D. (1992). The Outer Membrane , Not a Coat of Host Proteins , Limits Antigenicity of Virulent Treponema pallidum. Infect. Immun. *60*, 1076–1083.

Crooks, G.E., Hon, G., Chandonia, J., and Brenner, S.E. (2004). WebLogo : A Sequence Logo Generator. Genome Res. *14*, 1188–1190.

Cumberland, and Turner (1949). The rate of multiplication of Treponema pallidum in normal and immune rabbits. Am J Syph Gonorrhea Vener Dis. *33*, 201–2012.

Cunniffe, C., Ryan, F., Lambkin, H., and Brankin, B. (2012). Expression of tight and adherens junction proteins in cervical neoplasia. Br J Biomed Sci. *69*, 147–153.

Cunningham, T.M., Walker, E.M., Miller, J.N., and Lovett, M. a (1988). Selective release of the Treponema pallidum outer membrane and associated polypeptides with Triton X-114. J. Bacteriol. *170*, 5789–5796.

Dashper, S.G., Seers, C. a, Tan, K.H., and Reynolds, E.C. (2011). Virulence factors of the oral spirochete Treponema denticola. J. Dent. Res. *90*, 691–703.

Dayhoff, M., Schwartz, R., and Orcutt, B. (1978). A Model of Evolutionary Change in Proteins. Atlas Protein Seq. Struct. 345–352.

Deacon, W., Falcone, V., and Harris, A. (1957). A fluorescent test for treponemal antibodies. Proc Soc Exp Biol Med *96*, 477–480.

Dejana, E., and Orsenigo, F. (2013). Endothelial adherens junctions at a glance. J. Cell Sci. *126*, 2545–2549.

Dejana, E., Colella, S., Conforti, G., Abbadini, M., Gaboli, M., and Marchisio, P.C. (1988). Fibronectin and vitronectin regulate the organization of their respective Arg-Gly-Asp adhesion receptors in cultured human endothelial cells. J. Cell Biol. *107*, 1215–1223.

Drusin, L., Rouller, G., and Chapman, G. (1969). Electron Microscopy of Treponema pallidum Occurring. J. Bacteriol. *97*, 951–955.

Eddy, S.R. (1998). Profile hidden Markov models. Bionformatics Rev. *14*, 755–763.

Eddy, S.R. (2010). HMMER User ' s Guide. 1–93.

Evans, N.J., Brown, J.M., Demirkan, I., Murray, R.D., Birtles, R.J., Hart, C.A., and Carter, S.D. (2009). Treponema pedis sp. nov., a spirochaete isolated from bovine digital dermatitis lesions. Int. J. Syst. Evol. Microbiol. *59*, 987–991.

Felsenstein, J. (1985). Confidence Limits on Phylogenies : An Approach Using the Bootstrap. Evolution (N. Y). *39*, 783–791.

Fenno, J.C., Müller, K.H., and McBride, B.C. (1996). Sequence analysis, expression, and binding activity of recombinant major outer sheath protein (Msp) of Treponema denticola. J. Bacteriol. *178*, 2489–2497.

Fieldsteel, H., Cox, D.L., and Moeckli, R. (1981). Cultivation of virulent Treponema pallidum in tissue culture. Infect. Immun. *32*, 908–915.

Fieldsteel, H., Cox, D.L., and Moeckli, R. (1982). Further studies on replication of virulent Treponema pallidum in tissue cultures of Sf1Ep cells. Infect. Immun. *35*, 449–455.

Finlay, B.B., and Falkow, S. (1997). Common themes in microbial pathogenicity revisited. Microbiol. Mol. Biol. Rev. *61*, 136–169.

Finn, R.D., Clements, J., and Eddy, S.R. (2011). HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. *39*, 29–37.

Fitzgerald, T.J., and Repesh, L. a (1985). Interactions of fibronectin with Treponema pallidum. Genitourin. Med. *61*, 147–155.

Fitzgerald, T.J., Miller, J.N., and Sykes, J. a (1975). Treponema pallidum (Nichols strain) in tissue cultures: cellular attachment, entry, and survival. Infect. Immun. *11*, 1133–1140.

Fitzgerald, T.J., Johnson, R.C., Miller, J.N., and Sykes, J. a (1977a). Characterization of the attachment of Treponema pallidum (Nichols strain) to cultured mammalian cells and the potential relationship of attachment to pathogenicity. Infect. Immun. *18*, 467–478.

Fitzgerald, T.J., Cleveland, P., Johnson, R.C., Miller, J.N., and Sykes, J. a (1977b). Scanning electron microscopy of Treponema pallidum (Nichols strain) attached to cultured mammalian cells. J. Bacteriol. *130*, 1333–1344.

Fitzgerald, T.J., Repesh, L. a, Blanco, D.R., and Miller, J.N. (1984). Attachment of Treponema pallidum to fibronectin, laminin, collagen IV, and collagen I, and blockage of attachment by immune rabbit IgG. Br. J. Vener. Dis. *60*, 357–363.

Folkman, J., and Klagsbrun, M. (1987). Angiogenic Factors. Science (80-.). *235*, 442–447.

Frank, R.M. (1980). Bacterial penetration in the apical pocket wall of advanced human periodontitis. J. Periodontal Res. *15*, 563–573.

Fraser, C.M. (1998). Complete Genome Sequence of Treponema pallidum, the Syphilis Spirochete. Science (80-.). *281*, 375–388.

Gascuel, O., and Steel, M. (2006a). Neighbor-joining revealed. Mol. Biol. Evol. *23*, 1997–2000.

Gascuel, O., and Steel, M. (2006b). Neighbor-joining revealed. Mol. Biol. Evol. *23*, 1997–2000.

Ghanem, K.G., Moore, R.D., Rompalo, A.M., Erbelding, E.J., Jonathan, M., and Gebo, K.A. (2009). Lumbar puncture in HIV-infected patients with syphilis and no neurologic symptoms. Clin Infect Dis *48*, 816–821.

Gjestland, T. (1955). The Oslo study of untreated syphilis; an epidemiologic investigation of the natural course of the syphilitic infection based upon a re-study of the Boeck-Bruusgaard material. Acta Derm. Venereol. Suppl. (Stockh). *35*, 3 – 368.

Goldenberg, R. (2003). The infectious origins of stillbirth. Am. J. Obstet. Gynecol. *189*, 861–873.

Gomez, G.B., Kamb, M.L., Newman, L.M., Mark, J., Broutet, N., and Hawkes, S.J. (2013). Untreated maternal syphilis and adverse outcomes of pregnancy: a systematic review and meta-analysis. Bull. World Health Organ. *91*, 217–226.

Grab, D.J., Perides, G., Dumler, J.S., Kim, J., Park, J., Kim, Y. V, Nikolskaia, O., Choi, S., Stins, M.F., Kim, K.S., et al. (2005). Borrelia burgdorferi , Host-Derived Proteases , and the. Infect. Immun. *73*, 1014 – 1022.

Grab, D.J., Nyarko, E., Nikolskaia, O. V, Kim, Y. V, and Dumler, J.S. (2009). Human brain microvascular endothelial cell traversal by Borrelia burgdorferi requires calcium signaling. Clin. Microbiol. Infect. *15*, 422–426.

Grab, D.J., Nikolskaia, O., Kim, Y. V, Lonsdale-eccles, J.D., Ito, S., Hara, T., Fukuma, T., Nyarko, E., Kim, K.J., Stins, M.F., et al. (2013). WITH AN IN VITRO MODEL OF THE HUMAN INTERACTIONS AFRICAN TRYPANOSOME. J Parasitol *90*, 970–979.

Graber, J.R., Leadbetter, J.R., and Breznak, J.A. (2004). Description of Treponema azotonutricium the First Spirochetes Isolated from Termite Guts primitia sp . nov ., the First Spirochetes Isolated from Termite Guts. Appl. Environ. Microbiol. *70*, 1315–1320.

Grant, D.S., Tashiro, K., Segui-Real, B., Yamada, Y., Martin, G.R., and Kleinman, H.K. (1989). Two different laminin domains mediate the differentiation of human endothelial cells into capillary-like structures in vitro. Cell *58*, 933–943.

Grant, D.S., Kleinman, H.K., and Martin, G.R. (1990). The role of basement membranes in vascular development. Ann. N. Y. Acad. Sci. *588*, 61–72.

Haake, D.A., and Lovett, M.A. (1994a). Interjunctional Invasion of Endothelial Cell Monolayers. Methods Enzymol. *236*, 447–463.

Haake, D.A., and Lovett, M.A. (1994b). Treponema pallidum. Methods Enzymol. *236*, 447–463.

Haapasalo, M., Singh, U., McBride, B.C., and Uitto, V.J. (1991). Sulfhydryl-dependent attachment of Treponema denticola to laminin and other proteins. Infect. Immun. *59*, 4230–4237.

Hall, B.G. (2013). Building phylogenetic trees from molecular data with MEGA. Mol. Biol. Evol. *30*, 1229–1235.

Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Serius *41*, 95–98.

Han, C., Gronow, S., Teshima, H., Lapidus, A., Nolan, M., Lucas, S., Hammon, N., Deshpande, S., Cheng, J.-F., Zeytun, A., et al. (2011). Complete genome sequence of Treponema succinifaciens type strain (6091). Stand. Genomic Sci. *4*, 361–370.

Handsfield, H.H., Lukehart, S. a, Sell, S., Norris, S.J., and Holmes, K.K. (1983). Demonstration of Treponema pallidum in a cutaneous gumma by indirect immunofluorescence. Arch. Dermatol. *119*, 677–680.

Hansen, E.B., Pedersen, P.E., Schouls, L.M., Severin, E., and van Embden, J.D. (1985). Genetic characterization and partial sequence determination of a Treponema pallidum operon expressing two immunogenic membrane proteins in Escherichia coli. J. Bacteriol. *162*, 1227–1237.

Harris, T.O., Shelver, D.W., Bohnsack, J.F., and Rubens, C.E. (2003). A novel streptococcal surface protease promotes virulence , resistance to opsonophagocytosis , and cleavage of human fibrinogen. J. Clin. Invest. *111*, 61–70.

Hayes, N.S., Muse, K.E., Collier, a M., and Baseman, J.B. (1977). Parasitism by virulent Treponema pallidum of host cell surfaces. Infect. Immun. *17*, 174–186.

Hazlett, K.R.O., Cox, D.L., Decaffmeyer, M., Bennett, M.P., Desrosiers, D.C., Vake, C.J. La, Vake, M.E. La, Bourell, K.W., Robinson, E.J., Brasseur, R., et al. (2005). TP0453 , a Concealed Outer Membrane Protein of Treponema pallidum , Enhances Membrane Permeability. J. Bacteriol. *187*, 6499–6508.

Health, P. (2007). 2004 Canadian Sexually Transmitted Infections Surveillance Report. Canada Commun. Dis. Rep. *33S1*.

Henneke, P., and Golenbock, D.T. (2002). Innate immune recognition of lipopolysaccharide by endothelial cells. Crit Care Med *30*, 207–213.

Herbert, L.J., and Middleton, S.I. (2012). An estimate of syphilis incidence in Eastern Europe. J. Glob. Health *2*, 010402.

Hesketh, T., Zhu, W.X., and Zhou, X.D. (2010). Syphilis and social upheaval in China. N. Engl. J. Med. *363*, 1088–1089.

Heuner, K., Grosse, K., Schade, R., and Gobel, U.B. (2000). A flagellar gene cluster from the oral spirochaete Treponema maltophilum. Microbiology *146*, 497–507.

Ho, E.L., and Lukehart, S.A. (2011). Review series Syphilis : using modern approaches to understand an old disease. J. Clin. Investig. *121*, 4584–4592.

Holt, S.C. (1978). Anatomy and chemistry of spirochetes. Microbiol. Rev. *42*, 114–160.

Houston, S., Hof, R., Francescutti, T., Hawkes, A., Boulanger, M.J., and Cameron, C.E. (2011). Bifunctional role of the Treponema pallidum extracellular matrix binding adhesin Tp0751. Infect. Immun. *79*, 1386–1398.

Houston, S., Hof, R., Honeyman, L., Hassler, J., and Cameron, C.E. (2012). Activation and proteolytic activity of the Treponema pallidum metalloprotease, pallilysin. PLoS Pathog. *8*, e1002822.

Huber, P. (2009). Endothelial cell-cell junctions in vessel formation. J Soc Biol *203*, 119–123.

Huber, J.D., Egleton, R.D., and Thomas, P. (2001). Molecular physiology and pathophysiology of tight junctions in the blood – brain barrier. Trends Neurosci. *24*, 719–725.

Isaacs, R.D., and Radolf, J.D. (1990). Expression in Escherichia coli of the 37-kilodalton endoflagellar sheath protein of Treponema pallidum by use of the polymerase chain reaction and a T7 expression system. Infect. Immun. *58*, 2025–2034.

Izard, J., Renken, C., Hsieh, C.-E., Desrosiers, D.C., Dunham-Ems, S., La Vake, C., Gebhardt, L.L., Limberger, R.J., Cox, D.L., Marko, M., et al. (2009). Cryo-electron tomography elucidates the molecular architecture of Treponema pallidum, the syphilis spirochete. J. Bacteriol. *191*, 7566–7580.

Jepsen, O., Hougen, K., and Birch-Andersen, A. (1968). Electron microscopy of treponema pallidum Nichols. Acta Pathol Microbiol Scand. *74*, 241–258.

Jin, F., Brockmeier, U., and Otterbach, F. (2012). New Insight into the SDF-1 / CXCR4 Axis in a Breast Carcinoma Model : Hypoxia-Induced Endothelial SDF-1 and Tumor Cell CXCR4 Are Required for Tumor Cell Intravasation. Mol. Cancer Res. 1021–1031.

Jinga, V. V, Gafencu, a, Antohe, F., Constantinescu, E., Heltianu, C., Raicu, M., Manolescu, I., Hunziker, W., and Simionescu, M. (2000). Establishment of a pure vascular endothelial cell line from human placenta. Placenta *21*, 325–336.

Johnson, R.C., Ritzi, D.M., and Livermore, B.P. (1973). Outer envelope of virulent Treponema pallidum. Infect. Immun. *8*, 291–295.

Josenhans, C., and Suerbaum, S. (2002). The role of motility as a virulence factor in bacteria. Int. J. Med. Microbiol. *291*, 605–614.

Juanpere-Rodero, N., Martin-Ezquerra, G., Fernandez-Casado, A., Magan-Perea, L., Garcia-Alguacil, M. a, Barranco-Sanz, C., Serrano-Figueras, S., Pujol-Vallverdu, R.M., and Lloreta-Trull, J. (2013). Cell and tissue interactions of Treponema pallidum in primary and secondary syphilitic skin lesions: an ultrastructural study of serial sections. Ultrastruct. Pathol. *37*, 36–42.

Kampmeier, R. (1964). The late manifestations of syphilis: skeletal, visceral and cardiovascular. Med. Clin. N. Am *48*, 667–697.

Keenan, J., Day, T., Neal, S., Cook, B., Perez-Perez, G., Allardyce, R., and Bagshaw, P. (2000). A role for the bacterial outer membrane in the pathogenesis of Helicobacter pylori infection. FEMS Microbiol. Lett. *182*, 259–264.

Kubitschek, H.E. (1990). Cell volume increase in Escherichia coli after shifts to richer media. J. Bacteriol. *172*, 94–101.

Kubota, Y., Kleinman, H.K., Martin, G.R., and Lawley, T.J. (1988). Role of laminin and basement membrane in the morphological differentiation of human endothelial cells into capillary-like structures. J. Cell Biol. *107*, 1589–1598.

Lafond, R.E., and Lukehart, S.A. (2006). Biological Basis for Syphilis. Clin. Microbiol. Reveiws *19*, 29–49.

Lancet, T. (2011). Syphilis on the rise in the USA. Lancet *378*, 542.

Larkin, M. a, Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P. a, McWilliam, H., Valentin, F., Wallace, I.M., Wilm, a, Lopez, R., et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947–2948.

Lee, J., Choi, H., Jung, J., Lee, M., Lee, J., and Lee, K. (2003). Receptors for Treponema pallidum attachment to the Surface and Matric Proteins of Cultured Human Dermal Microvascular Endothelial Cells. Yonsei Med. J. *44*, 371–378.

Lee, S.-H., Kim, K.-K., Rhyu, I.-C., Koh, S., Lee, D.-S., and Choi, B.-K. (2006). Phenol/water extract of Treponema socranskii subsp. socranskii as an antagonist of Toll-like receptor 4 signalling. Microbiology *152*, 535–546.

Li, X., and Wang, X. (2000). Application of real-time polymerase chain reaction for the quantitation of interleukin-1β mRNA upregulation in brain ischemic tolerance. Brain Res. Protoc. *5*, 211–217.

Liu, J., Howell, J.K., Bradley, S.D., Zheng, Y., Zhou, Z.H., and Norris, S.J. (2010). Cellular Architecture of Treponema pallidum: Novel Flagellum, Periplasmic Cone, and Cell Envelope as Revealed by Cryo- Electron Tomography. J Mol Biol *403*, 546–561.

Lovett, M.A., David, A., Fogelmant, A.M., and Miller, J.N. (1988). Treponema pallidum invades intercellular junctions of endothelial cell monolayers. Proc. Natl. Acad. Sci. U. S. A. *85*, 3608–3612.

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., Abraham, J., Adair, T., Aggarwal, R., Ahn, S.Y., et al. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet *380*, 2095–2128.

Lukehart, S., Hook, E., Baker-Zander, S., Collier, A., Critchlow, C., and Handsfield, H. (1988). Invasion of the Central Nervous System by Treponema pallidum: Implications for Diagnosis and Treatment. Ann. Intern. Med. *109*, 855–862.

Luscombe, N.M., Greenbaum, D., and Gerstein, M. (2001). What is bioinformatics ? An introduction and overview. Yearb. Med. Inform. 83–100.

Lux, R., Miller, J.N., and Park, N. (2001). Motility and Chemotaxis in Tissue Penetration of Oral Epithelial Cell Layers by Treponema denticola. Infect. Immun. *69*, 6276–6283.

Maciag, T., Kadish, J., Wilkins, L., Stemerman, M.B., and Weinstein, R. (1982). Organizational behavior of human umbilical vein endothelial cells. J. Cell Biol. *94*, 511–520.

Madri, J., Pratt, B., and Tucker, A. (1988). Phenotypic Modulation of Endothelial Cells by Transforming Growth Factor-13 Depends upon the Composition and Organization of the Extracellular Matrix. J. Cell Biol. *106*, 1375–1384.

Mahoney, J., and Bryant, K. (1934). The time element in the penetration of the genital mucosa of the rabbit by the Treponema pallidum. Vener. Dis. Inf. *15*, 1–5.

Man, S., Ubogu, E.E., Williams, K. a, Tucky, B., Callahan, M.K., and Ransohoff, R.M. (2008). Human brain microvascular endothelial cells and umbilical vein endothelial cells differentially facilitate leukocyte recruitment and utilize chemokines for T cell migration. Clin. Dev. Immunol. *2008*, 384982.

Mantovani, A., Bussolino, F., and Dejna, E. (1992). Cytokine regulation of endothelial cell function. FASEB J. *6*, 2591–2599.

Marra, C., Baker-Zander, S. a, Hook, E.W., and Lukehart, S. a (1991). An experimental model of early central nervous system syphilis. J. Infect. Dis. *163*, 825–829.

Marra, C.M., Maxwell, C.L., Smith, S.L., Lukehart, S. a, Rompalo, A.M., Eaton, M., Stoner, B.P., Augenbraun, M., Barker, D.E., Corbett, J.J., et al. (2004). Cerebrospinal fluid abnormalities in patients with syphilis: association with clinical and laboratory features. J. Infect. Dis. *189*, 369–376.

Martìn-Padura, I., Lostaglio, S., Schneemann, M., Williams, L., Romano, M., Fruscella, P., Panzeri, C., Stoppacciaro, a, Ruco, L., Villa, a, et al. (1998). Junctional adhesion molecule, a novel member of the immunoglobulin superfamily that distributes at intercellular junctions and modulates monocyte transmigration. J. Cell Biol. *142*, 117–127.

Matsubara, K., Matsubara, Y., King, A., Zheng, J., Abe, E., Ito, M., and Maness, R. (2008). Estrogen-induced Angiogenesis and Co-factors. In Cell Growth Processes: New Research, D. Kimura, ed. (Nova Science Publishers), pp. 168–170.

Mengelle, C., Pasquier, C., Rostaing, L., Sandres-Sauné, K., Puel, J., Berges, L., Righi, L., Bouquies, C., and Izopet, J. (2003). Quantitation of human cytomegalovirus in recipients of solid organ transplants by real-time quantitative PCR and pp65 antigenemia. J. Med. Virol. *69*, 225–231.

Mocellin, S., Rossi, C.R., and Marincola, F.M. (2003). Quantitative real-time PCR in cancer research. Arch. Immunol. Ther. Exp. (Warsz). *51*, 301–313.

Morrison, D. (1996). Phylogenetic Tree-building. Int. J. Parasitol. *26*, 589–617.

Müller, M.M., and Griesmacher, a (2000). Markers of endothelial dysfunction. Clin. Chem. Lab. Med. *38*, 77–85.

Nassif, X., Bourdoulous, S., Eug, E., and Couraud, P. (2002). How do extracellular pathogens cross the blood – brain barrier ? Trends Microbiol. *10*, 227–232.

Norris, S.J. (1982). In vitro cultivation of Treponema pallidum: independent confirmation. Infect. Immun. *36*, 437–439.

Nusbaum, M.R.H., Wallace, R.R., Slatt, L.M., and Kondrad, E.C. (2004). Sexually transmitted infections and increased risk of co-infection with human immunodeficiency virus. J. Am. Osteopath. Assoc. *104*, 527–535.

Ohbayashi, T., Irie, A., Murakami, Y., Nowak, M., Potempa, J., Nishimura, Y., Shinohara, M., and Imamura, T. (2011). Degradation of fibrinogen and collagen by staphopains, cysteine proteases released from Staphylococcus aureus. Microbiology *157*, 786–792.

Ohta, K., Makinen, K.K., and Loesche, W.J. (1986). Purification and characterization of an enzyme produced by Treponema denticola capable of hydrolyzing synthetic trypsin substrates. Infect. Immun. *53*, 213–220.

Olsen, I. (1984). Attachment of Treponema denticola to cultured human epithelial cells. Scand. J. Dent. Res. *92*, 55–63.

Park, K.-K., Heuner, K., Göbel, U.B., Yoo, Y.-J., Kim, C.-K., and Choi, B.-K. (2002). Cloning and characterization of a major surface protein (MspTL) of Treponema lecithinolyticum associated with rapidly progressive periodontitis. FEMS Microbiol. Lett. *207*, 185–192.

Paster, B., and Dewhirst, F. (2000). Phylogenetic foundation of spirochetes. J Mol Microbiol Biotechnol *2*, 341–344.

Patti, J.M., Allen, B.L., McGavin, M.J., and Höök, M. (1994). MSCRAMM-mediated adherence of microorganisms to host tissues. Annu. Rev. Microbiol. *48*, 585–617.

Penn, C.W., Cockayne, a, and Bailey, M.J. (1985). The outer membrane of Treponema pallidum: biological significance and biochemical properties. J. Gen. Microbiol. *131*, 2349–2357.

Peters, S.R., Valdez, M., Riviere, G., and Thomas, D.D. (1999). Adherence to and penetration through endothelial cells by oral treponemes. Oral Microbiol. Immunol. *14*, 379–383.

Polanco, J.C., Rodríguez, J.A., Corredor, V., and Patarroyo, M.A. (2002). Plasmodium vivax: parasitemia determination by real-time quantitative PCR in Aotus monkeys. Exp. Parasitol. *100*, 131–134.

Ponchel, F., Toomes, C., Bransfield, K., Leong, F.T., Douglas, S.H., Field, S.L., Bell, S.M., Combaret, V., Puisieux, A., Mighell, A.J., et al. (2003). Real-time PCR based on SYBR-Green I fluorescence: an alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. BMC Biotechnol. *3*, 1–13.

Public Health Agency of Canada (2008). Report on Sexually Transmitted Infections in Canada : 2008.

Quist, E.E., Repesh, L. a, Zeleznikar, R., and Fitzgerald, T.J. (1983). Interaction of Treponema pallidum with isolated rabbit capillary tissues. Br. J. Vener. Dis. *59*, 11–20.

Radolf, J.D. (1995). MicroReview Treponema pallidum and the quest for outer membrane proteins. Mol. Microbiol. *16*, 1067–1073.

Radolf, J.D., and Norgard, M. V (1988). Pathogen specificity of Treponema pallidum subsp. pallidum integral membrane proteins identified by phase partitioning with Triton X-114. Infect. Immun. *56*, 1825–1828.

Radolf, J.D., Norgard, M. V, and Schulz, W.W. (1989). Outer membrane ultrastructure explains the limited antigenicity of virulent Treponema pallidum. Proc. Natl. Acad. Sci. U. S. A. *86*, 2051–2055.

Radolf, J.D., Robinson, E.J., Bourell, K.W., Akins, D.R., Porcella, S.F., Weigel, L.M., Jones, J.D., and Norgard, M. V (1995). Characterization of outer membranes isolated from Treponema pallidum, the syphilis spirochete. Infect. Immun. *63*, 4244–4252.

Riley, B., Oppenheimer-Marks, N., Hansen, E., Radolf, J., and Norgard, M. (1992). Virulent Treponema pallidum activates human vascular endothelial cells. J Infect Dis. *165*, 484–493.

Riviere, G.R., Thomas, D.D., and Cobb, C.M. (1989). In vitro model of Treponema pallidum invasiveness. Infect. Immun. *57*, 2267–2271.

Riviere, G.R., Weisz, K.S., Adams, D.F., and Thomas, D.D. (1991). Pathogen-related oral spirochetes from dental plaque are invasive. Infect. Immun. *59*, 3377–3380.

Rose'Meyer, R.B., Mellick, A.S., Garnham, B.G., Harrison, G.J., Massa, H.M., and Griffiths, L.R. (2003). The measurement of adenosine and estrogen receptor expression in rat brains following ovariectomy using quantitative PCR analysis. Brain Res. Brain Res. Protoc. *11*, 9–18.

Salazar, J.C., Hazlett, K.R.O., and Radolf, J.D. (2002). The immune response to infection with Treponema pallidum, the stealth pathogen. Microbes Infect. *4*, 1133–1140.

Salazar, J.C., Cruz, A.R., Pope, C.D., Valderrama, L., Saravia, N.G., and Radolf, J.D. (2007). Treponema pallidum Elicits Innate and Adaptive Cellular Immune Responses in Skin and Blood during Secondary Syphilis: A Flow- Cytometric Analysis. J Infect Dis *195*, 879–887.

Schneeberger, E.E., Walters, D. V, and Olver, R.E. (1978). Development of intercellular junctions in the pulmonary epithelium of the foetal lamb. J. Cell Sci. *32*, 307–324.

Schneider, T.D., and Stephens, R.M. (1990). sequences. Nucleic Acids Res. *18*, 6097–6100.

Schwan, T.G., Simpson, W.J., Schrumpf, M.E., and Karstens, R.H. (1989). Identification of Borrelia burgdorferi and B. hermsii using DNA hybridization probes. J. Clin. Microbiol. *27*, 1734–1738.

Sell, S., Gamboa, D., Baker-Zander, S., Lukehart, S., and Miller, J. (1980). Host Response to Treponema pallidum in Intradermally-infected Rabbits: Evidence for Persistence of Infection at Local and Distant SItes. J. Invest. Dermatol. *75*, 470–475.

Sheffield, J.S., Wendel, G.D., McIntire, D.D., and Norgard, M. V (2007). Effect of genital ulcer disease on HIV-1 coreceptor expression in the female genital tract. J. Infect. Dis. *196*, 1509–1516.

Sigma-Aldrich (2011). Cell Viability and Proliferation. Biofiles *6*, 17–22.

Sigma-Aldrich (2013). Lactate Dehydrogenase Activity Assay Product Sheet. 1–4.

Simms, I., Fenton, K. a., Ashton, M., Turner, K.M.E., Crawley-Boevey, E.E., Gorton, R., Thomas, D.R., Lynch, A., Winter, A., Fisher, M.J., et al. (2005). The Re-Emergence of Syphilis in the United Kingdom: The New Epidemic Phases. Sex. Transm. Dis. *32*, 220–226.

Singh, Ameeta Eshri, Karen Sutherland, Bonita Lee, Joan Louise Robinson, T.W. (2007). Public Health Resurgence of early congenital syphilis in Alberta. Practice *177*, 3–6.

Slivienski-gebhardt, L.L., Izard, J., Samsonoff, W.A., and Limberger, R.J. (2004). Development of a Novel Chloramphenicol Resistance Expression Plasmid Used for Genetic Complementation of a fliG Deletion Mutant in Treponema denticola. Infect. Immun. *72*, 5493–5497.

Van der Sluis, J.J., Koehorst, J. a, and Boer, a M. (1987). Factors that inhibit adherence of Treponema pallidum (Nichols strain) to a human fibroblastic cell line: development in serum of patients with syphilis. Genitourin. Med. *63*, 71–76.

Soltis, D.E., and Soltis, P.S. (2003). Applying the Bootstrap in Phylogeny Reconstruction. Stat. Sci. *18*, 256–267.

Stamm, L. V, and Bassford, P.J. (1985). Cellular and extracellular protein antigens of Treponema pallidum synthesized during in vitro incubation of freshly extracted organisms. Infect. Immun. *47*, 799–807.

Stokes, J., Beerman, H., and Ingraham, N. (1944). Modern clinical syphilology.

Swain, R. (1955). Electron microscopic studies of the morphology of pathogenic spirochaetes. J Pathol Bacteriol. *69*, 117–128.

Swancutt, M. a, Riley, B.S., Radolf, J.D., and Norgard, M. V (1989). Molecular characterization of the pathogen-specific, 34-kilodalton membrane immunogen of Treponema pallidum. Infect. Immun. *57*, 3314–3323.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. *24*, 1596–1599.

Tantalo, L.C., Lukehart, S.A., Marra, C.M., The, S., Diseases, I., and Jan, N. (2005). Treponema pallidum Strain-Specific Differences in Neuroinvasion and Clinical Phenotype in a Rabbit Model. J. Infect. Dis. *191*, 75–80.

Tchaptchet, S., Gumenscheimer, M., Kalis, C., Freudenberg, N., Hölscher, C., Kirschning, C.J., Lamers, M., Galanos, C., and Freudenberg, M. a (2012). TLR9-dependent and independent pathways drive activation of the immune system by Propionibacterium acnes. PLoS One *7*, e39155.

Thomas, D.D., and Higbie, L.M. (1990). In vitro association of leptospires with host cells. Infect. Immun. *58*, 581–585.

Thomas, D., Fogelman, A., Miller, J., and Lovett, M. (1989). Interactions of Treponema pallidum with Endothelial Cell Monolayers. Eur. J. Epidemiol. *5*, 15–21.

Trott, D.J., Moeller, M.R., Zuerner, R.L., Goff, J.P., Waters, W.R., Alt, D.P., Richard, L., Wannemuehler, M.J., and Walker, R.L. (2003). Characterization of Treponema phagedenis -Like Spirochetes Isolated from Papillomatous Digital Dermatitis Lesions in Dairy Cattle Characterization of Treponema phagedenis -Like Spirochetes Isolated from Papillomatous Digital Dermatitis Lesions in Dairy Ca. J. Clin. Microbiol. *41*, 2522–2529.

Turner, T. (1957). Biology of the treponematoses. World Heal. Organ. *35*, 3–266.

Vergnolle, N., Derian, C.K., D'Andrea, M.R., Steinhoff, M., and Andrade-Gordon, P. (2002). Characterization of Thrombin-Induced Leukocyte Rolling and Adherence: A Potential Proinflammatory Role for Proteinase-Activated Receptor-4. J. Immunol. *169*, 1467–1473.

Vestweber, D. (2000). Molecular mechanisms that control endothelial cell contacts. J. Pathol. *190*, 281–291.

Walker, J.M. (2009). Angiogenesis Protocols (New Yorl: Springer Protocols. Humana Pres).

Walker, E.M., Zampighi, G. a, Blanco, D.R., Miller, J.N., and Lovett, M. a (1989). Demonstration of rare protein in the outer membrane of Treponema pallidum subsp. pallidum by freeze-fracture analysis. J. Bacteriol. *171*, 5005–5011.

Walker, E.M., Borenstein, L. a, Blanco, D.R., Miller, J.N., and Lovett, M. a (1991). Analysis of outer membrane ultrastructure of pathogenic Treponema and Borrelia species by freeze-fracture electron microscopy. J. Bacteriol. *173*, 5585–5588.

Wallace, A., and Harris, A. (1967). Reiter treponeme: a review of the literature. Bull. W. H. O. *36*, 5–7.

Watson-Jones, D., Changalucha, J., Gumodoka, B., Weiss, H., Rusizoka, M., Ndeki, L., Whitehouse, A., Balira, R., Todd, J., Ngeleja, D., et al. (2002). Syphilis in pregnancy in Tanzania. I. Impact of maternal syphilis on outcome of pregnancy. J. Infect. Dis. *186*, 940–947.

Weiser, J.N., and Gotschlich, E.C. (1991). Outer membrane protein A (OmpA) contributes to serum resistance and pathogenicity of Escherichia coli K-1. Infect. Immun. *59*, 2252–2258.

Whitman, W. (2010). Bergey's Manual of Systematic Bacteriology.

Wilson-welder, J.H., Elliott, M.K., Zuerner, R.L., Bayles, D.O., and Alt, D.P. (2013). Biochemical and molecular characterization of Treponema phagedenis-like spirochetes isolated from a bovine digital dermatitis lesion. BMC Microbiol. *13*.

World Health Organization (2007). GLOBAL STRATEGY FOR THE PREVENTION AND CONTROL OF SEXUALLY TRANSMITTED INFECTIONS: 2006–2015.

World Health Organization (2011). Global HIV/AIDS Response.Epidemic update and health sector progress towards Universal Access.

World Health Organization (2012). Global incidence and prevalence of selected curable sexually transmitted infections - 2008. WHO Press.

Yun, J.J., Heisler, L.E., Hwang, I.I.L., Wilkins, O., Lau, S.K., Hyrcza, M., Jayabalasingham, B., Jin, J., McLaurin, J., Tsao, M.-S., et al. (2006). Genomic DNA functions as a universal external standard in quantitative real-time PCR. Nucleic Acids Res. *34*, e85.

Zoechling, N., Schluepen, E.M., Soyer, H.P., Kerl, H., and Volkenandt, M. (1997). Molecular detection of Treponema pallidum in secondary and tertiary syphilis. Br. J. Dermatol. *136*, 683–686.