Resolution and characterization of subgroups of *Gardnerella vaginalis* and description of the vaginal microbiota of women with preterm premature rupture of membranes

Submitted to the College of Graduate Studies and Research of the University of Saskatchewan in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Veterinary Microbiology at the University of Saskatchewan.

By

TEENUS PARAMEL JAYAPRAKASH

PERMISSION TO USE

In agreement with the outlines set out by the College of Graduate Studies and Research at the University of Saskatchewan, I allow the University of Saskatchewan Libraries to make this thesis available to all interested parties. Also in accordance with the College of Graduate Studies and Research, I allow this thesis to be copied "in any manner, in whole or in part, for scholarly purposes". This thesis may not, however, be reproduced or used in any manner for financial gain without my written consent. Any scholarly use of this thesis, in part or in whole, must acknowledge both myself and the University of Saskatchewan.

Any requests for copying or using this thesis, in any form or capacity, should be made to:

Head of Department of Veterinary Microbiology

University of Saskatchewan

Saskatoon, Saskatchewan

S7N 5B4

ABSTRACT

The vaginal microbial community is critical to a woman's health and the health of her family. Bacterial vaginosis (BV) is a polymicrobial syndrome characterized by a shift of the vaginal microbiota from a *Lactobacillus* dominated community to a dense biofilm containing a complex mixture of organisms. Although BV is an important risk factor for poor reproductive health outcomes, the etiology of BV is poorly understood.

Gardnerella vaginalis is a hallmark species of BV. Phylogenetic analysis of *cpn*60 universal target sequences from metagenomic studies of the vaginal microbiome and from *G. vaginalis* isolates resolved four subgroups within the species. This subdivision, supported by whole genome similarity comparisons, demonstrated that these subgroups might represent different species. Among a group of African women, only *G. vaginalis* subgroup B was significantly more abundant in women with BV relative to women with Nugent scores not consistent with BV. To characterize the subgroups further, several phenotypic and molecular factors of *G. vaginalis* subgroups were assessed. Proteomic profiles of isolates within each subgroup formed unambiguous clusters. Sialidase gene sequences were detected in all subgroups, however enzymatic activity was detected only in subgroup B. Two isolates of subgroup B isolates (N153 and N101) were incapable of growth in 7% CO_2. Given the well-known relationship between an anaerobic microbiota and BV, anaerobic isolates of *G. vaginalis* are potentially important players in the vaginal microbial community. To determine genome content differences that could account for the phenotypic difference, whole genome sequences of four *G. vaginalis* subgroup B isolates representing facultative and anaerobic phenotypes were determined. Comparison of genomes led to the identification of genes predicted to

encode proteins involved in cell wall biogenesis and protection from oxidative damage that might account for the observed phenotypes.

The *cpn*60 universal target based methodology that improved resolution of the vaginal microbiota including *G. vaginalis* was applied in a prospective study of the vaginal microbiome of women with preterm premature rupture of membranes (PPROM). The objectives were to characterize the vaginal microbiota of women following PPROM, and to determine if microbiome composition at the time of rupture predicts latency duration and perinatal outcomes. Only 13/70 samples collected from 36 women were dominated by *Lactobacillus* spp., the expected profile for healthy women, while *Megasphaera* type 1 and *Prevotella* spp. were detected in all samples. Microbiome profiles at the time of membrane rupture did not cluster by gestational age at PPROM, or latency duration. Microbial profiles were unstable over the latency period, with dramatic shifts in composition between weekly samples, and an overall decrease in *Lactobacillus* abundance. Mollicutes were detected by PCR in 81% (29/36) of women, and these women had significantly lower gestational age at delivery and correspondingly lower birth weight infants than Mollicutes negative women.

Taken together, the results presented in this thesis demonstrate the value of high resolution profiling of the vaginal microbiome using *cpn*60 UT sequences. The resolution of subgroups within *G. vaginalis* has potentially significant implications for women's health diagnostics, requiring a shift away from considering *G. vaginalis* as a single entity. The PPROM study provides foundational information that may lead to the identification of informative sequence patterns, providing clinicians with better tools for expectant management following PPROM.

ACKNOWLEDGEMENTS

I owe my thesis and my PhD degree to many people who believed in me and thought I could do what I did.

A great scientist, a true professor, an amazing human being, a great listener, one of the best supervisors on this planet - Dr. Janet hill. I am grateful to you for choosing me to be a member of the Hill Lab family. I want you to know that you are one of the best people life gifted me in my life. I am indebted to Dr. Janet for the patience she rendered while teaching me science. She taught me to think. She believed in me when I could not. She taught me values of determination, discipline and the importance of being there before time. Your teachings will truly serve as my guiding light all the way till true darkness.

I wish to thank my committee, who build my confidence over the years. I wish to extend my special thanks to Dr. Lessard for introducing me to the world of proteins.

I have been lucky to be part of the PPROM team, under the leadership of Dr. Deborah money. Thank you Emily for answering my questions every time with so much enthusiasm. I am so thankful to the PPROM research team (Arianne, Daljeet, Sandy, Fontayne, Shivinder, Nancy, Amber, Zoye, Melissa) in BC Women's Hospital in Vancouver for helping me with the PPROM research and for collecting samples from those mothers who were having the most stressful time of their lives. Thank you mothers.

"You don't choose your family. They are God's gift to you, as you are to them" ~Desmond Tutu. I am lucky I found you – Champika, Bonnie, Isha, Matt, Atul, John. I learnt all that I know to do in wet lab from you both – Champika and Bonnie. Thank you Champika for being our mom in the lab. Thank you Kosala Rajapaksha for teaching me how to run SDS PAGE. My office mates, without you – Rui, Iran, Yanyun, Shrijana and all other officemates – life would not have been so enlightening and entertaining as it has been for the past five years. Thank you Lana, Linda and Pat.

I wish to take this opportunity to thank Sir and chechi (Suresh and Viji Kalagnanam) for giving me an unofficial surname in Saskatoon – Kalagnanam. Thank you for adopting me into your lives.

Thank you Amma (Mom) for being the first and the most important teacher of my life. Thank you Acha (Dad) for convincing me to dream and to go for it. Thank you Amma, my mother-in-law for asking her son to support me. Thank you Wisson, my little brother for letting me be that sister.

Finally thanks to my husband, who did more than I have done for anybody, who loved me enough to come for me to a world unknown. And to my son (Maanav) who waited in the world unknown for 41 week to let me finish my thesis and who slept well the night before my defense.

To the power unknown that lets me do what I need to do.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

CHAPTER 1 - Introduction and literature review

1.1 The vaginal microbiome

The vaginal microbiome, a complex microbial community residing in the vagina, is a dynamic ecosystem susceptible to changes in structure and composition of species present. The delicate equilibrium that exists between vaginal epithelial cells and the microbial community can be disturbed, resulting in an unbalanced microbial community. Such imbalances in the vaginal microbial community can cause conditions such as bacterial vaginosis (BV) (Spiegel 1991), yeast infections (Egan and Lipsky 2000) and pelvic inflammatory disease (PID) (Ness *et al.* 2004). A disturbed microbial community also increases susceptibility to sexually transmitted diseases (Myer *et al.* 2005) such as chlamydia (caused by *Chlamydia trachomatis*), trichomoniasis (caused by *Trichomonas vaginalis*) and acquired immunodeficiency syndrome (AIDS, caused by HIV). Perturbed vaginal microbiota is also associated with preterm birth and preterm rupture of membranes (Leitich *et al.* 2003).

A healthy vaginal ecosystem has been traditionally defined by the dominance of certain *Lactobacillus* species, which suppress the overgrowth of other endogenous bacteria by the production of lactic acid, H_2O_2, and bacteriocins (Dover *et al.* 2008; O'Hanlon *et al.* 2013; Summers 2010). However, the vaginal microbial community is a dynamic system, with different species attaining dominance throughout a woman's lifetime under the influence of variety of host factors such as hormonal changes (Eschenbach *et al.* 2000), vaginal epithelial changes (Sjoberg *et al.* 1988), metabolites available (Huggins and Preti 1976), the host's genetic make up and mechanisms of innate immunity (Cole 2006).

Glycogen metabolism in the vagina is known to be critical for the maintenance of an acidic pH. Cruickshank and Sharman in 1934 recorded for the first time the presence of glycogen in the fetal vagina and also in the infants up until the fourth week of life (reviewed by Henderson *et al.* (1969)). The infant vagina at birth has been shown to contain residues of transplacental maternal estrogen. Estrogen causes proliferation of the glycogen rich intermediate cell layer of the mucosa, which in turn causes the vaginal mucosa of infants to be rich in glycogen. Hence, glycogen in the infant vagina can support the growth of lactic acid producing microbes such as lactobacilli. Lactobacilli are detected in vagina of infants up to 48 hours after birth. Schroder, Hinrichs and Kessler in 1926 had proposed that vaginal glycogen was metabolized into simpler carbohydrates, which were later metabolized by the lactobacilli (reviewed by Wylie and Henderson (1969)).

Maternal estrogen wanes by the fourth post-natal week causing the infant vaginal epithelium to thin out, also losing its glycogen content. During early childhood (1-8 years), the vaginal pH is either neutral or alkaline due to lack of acid producing microbes in the vagina, associated with the lack of glycogen. There is hardly any stimulation by adrenal or gonadal steroid hormones in the vagina until puberty (Farage and Maibach 2006; Farage and Maibach 2011). During the childhood years, the vaginal microbial community is composed primarily of enteric microbes and also resembles skin microbiota with diphtheroids (corynebacteria) and *Staphylococcus epidermidis* being prominent members of the community (Hammerschlag *et al.* 1978).

As puberty begins (between 8 and 13 years of age), adrenal and gonadal glands mature, the vaginal epithelium thickens and stratifies, and the production of intracellular

glycogen in the vagina ensues (Hammerschlag *et al.* 1978). Cervicovaginal secretions are also produced and prevalence of lactic acid producing microbes rises in the vagina (Farage and Maibach 2006).

During the reproductive years, with menarche approaching and the menstrual cycle inducing cyclic hormonal changes, estrogen levels rise and vaginal epithelial thickness increases (Summers 2010), also leading to parakeratosis (Nauth and Haas 1985). Glycogen content in the vagina rises, bringing about the dominance of lactic acid producing microbes (Eschenbach *et al.* 2000). Lactic acid, a byproduct produced by vaginal microbiota as well as by vaginal mucosal metabolism, establishes the pH of vagina to be acidic, as low as 3.5 (Boskey *et al.* 2001; O'Hanlon *et al.* 2013), which is indicative of a healthy vaginal environment. Over the course of a menstrual cycle, growth of *Lactobacillus* increases with recovery of non-*Lactobacillus* spp. being higher only during menses, according to the results of culture based studies (Bartlett *et al.* 1977; Eschenbach *et al.* 2000; Onderdonk *et al.* 1986). However, results of a more recent culture independent study by Chaban *et al.,*(2014) contradicts this observation of reduced *Lactobacillus* during menses. Another culture independent longitudinal study of two consecutive menstrual cycles report reduction in *L. crispatus* during menses and no change in any other *Lactobacillus* species (Santiago *et al.* 2011a).

The most commonly detected species of vaginal *Lactobacillus* are *L. iners, L. crispatus, L. gasseri*, and *L. jensenii* (Ravel *et al.* 2011). During the reproductive years, colonization by 'undesirable' organisms also occurs and these non-*Lactobacillus* spp. are often associated with unhealthy states such as BV, yeast infection, sexually transmitted diseases and urinary tract infections. These 'undesirable' taxa that are often isolated from

women with abnormal vaginal microbiota include *Gardnerella, Staphylococcus, Ureaplasma, Corynebacterium, Streptococcus, Peptostreptococcus, Mycoplasma, Enterococcus, Escherichia, Veillonella, Bifidobacterium* and *Candida* (Chaban *et al.* 2014; Eschenbach *et al.* 2000; Ravel *et al.* 2011; Zhou *et al.* 2004).

Pregnancy brings about a temporary disruption in the physiology during the reproductive years that causes many changes including an increase in blood volume, specifically to the vulvar region and relaxation of the connective tissue in the body. The menstrual cycle ceases and progesterone and estrogen levels are elevated during gestation (Magness 1998). The risk of *Candida* infection is documented to increase during pregnancy (Soong and Einarson 2009; Wallenburg and Wladimiroff 1976). *Lactobacillus* spp. are described to be the predominant members of the microbial community in normal pregnancy, with 'normal' being defined by the Romero *et al.* as one without obstetrical, medical or surgical complications, and delivery occurring at term (38 to 42 weeks) (Romero *et al.* 2014a). The stability of the vaginal microbial community has been reported to be higher in pregnant women than in non-pregnant women and the relative abundance of taxa associated with BV such as *Atopobium, Prevotella*, *Sneathia, Gardnerella,* Ruminococcaceae*, Parvimonas, Mobiluncus* was found to be lower in pregnant women in comparison to non-pregnant women (Romero *et al.* 2014b).

During menopause, the follicular function as well as the menstrual cycle ceases (Farage and Maibach 2006). The vaginal epithelium atrophies and becomes dry, with a distinct reduction in estrogen production in the vagina (Pandit and Ouslander 1997). Glycogen content in the vagina decreases, becoming insufficient to support the growth of *Lactobacillus* and other lactic acid producing microbes, which leads to a corresponding

rise in vaginal pH. Cervicovaginal secretions become sparse, causing a decline in protective mucosal secretions, increasing the likelihood of urinary tract infections and vaginal yeast infections (Farage and Maibach 2006; Summers 2010). Atrophic vaginitis is also common in postmenopausal women (with low relative abundance of *Lactobacillus*) (Brotman *et al.* 2014; Pandit and Ouslander 1997). Lactobacilli isolated from postmenopausal women such as *L. casei* are more of gut origin than species associated with the vagina of reproductive aged women (Petricevic *et al.* 2013). Colonization by enteric organisms increases and microbes detected in reproductive years such as *G. vaginalis, Ureaplasma urealyticum, Candida albicans* are less frequently isolated from postmenopausal women (Hillier and Lau 1997).

1.2 Methodology for studying vaginal microbiota

Until a decade ago, methods of studying the vaginal microbial community were mainly culture based. Culture-based methods have been used for decades for describing microbes inhabiting various sites of the human body, including the genital tract of women. Culture-based methods have contributed significantly to our current understanding of the vaginal microbial community, describing the microbes associated with health and disease, and also the physiological and metabolic potential of these microbes to a great extent. The concept of a lactobacilli dominated microbiota being associated with a healthy genital tract in women is one of the earliest observations based on this approach. Understanding of the ability of *Lactobacillus* to produce lactic acid, hydrogen peroxide and bacteriocins suppressing the growth of undesirable microbes in vagina was established early on using culture based studies. Similarly, culture-based

studies led to the identification of *G. vaginalis, Mycoplasma hominis, Bacteriodes, Prevotella* and *Mobiluncus* as taxa associated with the most common vaginal disorder, BV (Spiegel 1991).

However, culture based methods do have their disadvantages such as being laborious and time consuming. Culture based methods also lead to the isolation of only plate-growth-adapted organisms. Many organisms are never isolated due to their special nutrient requirements. This in turn leads to a significant underestimation of the diversity and richness of the vaginal microbial community of even healthy women. In the face of these limitations, culture independent methods provide a more unbiased and complete assessment of the vaginal microbial community in healthy and unhealthy states.

With the introduction of culture independent methods based on DNA fingerprinting such as Denaturing Gradient Gel Electrophoresis (DGGE) and Terminal Restriction Fragment Length Polymorphism (TRFLP), many bacterial species were described as members of the vaginal microbiota and new taxa were associated with health and disease (Thies *et al.* 2007). DGGE based studies resulted in the discovery of *Lactobacillus iners* as an important member of vaginal microbial community (Martin *et al.* 2012). *L. iners* only grows on blood agar and hence was not considered as important as other vaginal lactobacilli, which unlike *L. iners* grow well on de Man, Rogosa, and Sharpe (MRS) agar (Vasquez *et al.* 2002). With research efforts being steered in the direction of sequencing based methods, cloning and single gene based sequencing started providing a more complicated picture of vaginal microbiota. *cpn*60 and 16S rRNA became the two most popular gene targets for metagenomic studies. Using these methods, significant microbial diversity was reported even within a healthy vaginal microbiota.

One of the earlier studies based on cloning and sequencing of 16S rRNA libraries described many species for the first time as being part of the vaginal microbial community such as *Megasphaera* and *Leptotrichia* (Zhou *et al.* 2004). Similarly the earliest study of *cpn*60 universal target (UT) based sequencing by Hill *et al.* (2005a) led to a high resolution picture of the vaginal microbiota of healthy non-pregnant Canadian women. Hill *et al.* reported for the first time *C. psittaci* and two other species of the family Chlamydiaceae as part of human vaginal microbiota (Hill *et al.* 2005a). The advantage of the *cpn*60 based method over the 16S rRNA based sequencing has been its ability to resolve sequences into phylogenetic groups as shown in case of the sequences from the family Chlamydiaceae (Hill *et al.* 2005a).

1.3 Bacterial vaginosis

Bacterial vaginosis, historically known as 'nonspecific vaginitis', has undergone several name changes since mid-1950s. In 1955, when Gardner and Dukes proposed *Gardnerella vaginalis* (then known as *Haemophilis vaginalis*) as the etiologic agent of nonspecific vaginitis, the name for non-specific vaginitis was changed to '*Haemophilis vaginalis* vaginitis' (Gardner and Dukes 1955). With the recognition that there is no inflammation present in the vagina with the abnormal discharge during the condition, 'vaginosis' was proposed to be the new name for the syndrome (Holmes *et al.* 1981). However, with increasing knowledge of many different bacterial species being associated with the syndrome, term 'bacterial vaginosis' was adopted and has been used since then.

BV is the most common infection for which women seek medical attention and an estimated 300 million women worldwide are thought to experience BV each year (Powell

2013). Generally, BV is characterized by a shift in the normal microbiome of the vagina from the predominant *Lactobacillus* species to an overgrowth of a mixed community dominated by *G. vaginalis* and anaerobic bacterial species. BV is diagnosed using mainly two methods: a microbiologic criterion using the Nugent score, and by specific clinical criteria described as Amsel's criteria. The Nugent score is calculated by observing a Gram stain of a vaginal smear. The Nugent score is based on the number of large Gram-positive rods (*Lactobacillus* morphotypes), and small Gram-variable rods (*G. vaginalis* morphotypes and *Mobiluncus* morphotypes). The score given for the Gram stain of the smear ranges from 0 to 10 (Nugent *et al.* 1991) with 0 - 3 being not consistent with BV (normal), 4 - 6 being intermediate and 7 - 10 being consistent with BV. Gram stain Nugent scoring is the current gold standard method for diagnosing BV due to its reproducibility in the hands of an expert technical staff scoring the Gram stain slide and also due to its sensitivity. However, its disadvantages include the requirement of a technical staff with years of experience and also requirements of laboratory facilities. Other major limitations of Nugent score are that it is quite low resolution and is only semi-quantitative. The Nugent score, being based on Gram staining, does not give specific information about what bacterial species are present.

Using Amsel's criteria (Amsel *et al.* 1983), a positive clinical diagnosis of BV is made when at least three of the four following criteria are present: thin homogeneous vaginal discharge, vaginal pH of ≥ 4.5, positive whiff test (fishy odor on addition of 10% KOH to the discharge), or presence of clue cells in a wet mount preparation. When compared to Nugent score as the gold standard test, diagnosis of BV done by Amsel's criteria has a clinical sensitivity of 97% and specificity of 90% (Simoes *et al.* 2006).

BV can also be defined based on the phylogenetic characteristics of a sample determined by sequences obtained using deep sequencing, which is referred to as a 'molecular definition' of BV. For instance, Schellenberg *et al*. (2011b) defined samples as 'BV negative' as ones with high abundance of *Lactobacillus* and 'BV positive' samples as those with lowest abundance of *Lactobacillus.* Samples defined as 'intermediate' were dominated by *Gardnerella*. Schellenberg *et al*. (2011b) also defined an outlier group dominated by *E. coli.* The presence of specific sequences such as that of BV associated bacteria (BVAB1 closely related to *Eubacterium* spp., BVAB2 and BVAB3 closely related to *Clostridium stercorarium*) has been strongly associated with the diagnosis of BV (Fredricks *et al.* 2005; Haggerty *et al.* 2009). *Prevotella* spp., BVAB1 and *Dialister micraerophilus* have been related to the clinical criterion of positive whiff test (amine odor upon addition of 10% KOH to vaginal discharge) (Srinivasan *et al.* 2012). Another important bacterial species described as a strong predictor of BV by culture independent studies is *Megasphaera* type 1 (Fredricks *et al.* 2007; Shipitsyna *et al.* 2013). Though these species are described in the literature as being strong predictors of BV, none of them are currently being used in any assays for diagnosis of BV.

BV in many women is often a recurrent or chronic condition. As already described, BV is mostly characterized by an overgrowth of predominantly anaerobic organisms like *G. vaginalis, Prevotella* spp*., Peptostreptococcus, Atopobium* and *Mobiluncus* spp. in the vagina leading to a replacement of lactobacilli. BV-associated organisms produce many virulence factors: mucinases, cytolysins, prolidases, siderophores, human-lactoferrin binding proteins, lipases, phospholipases and proteolytic

carboxylase enzymes. BV-associated organisms are capable of breaking down vaginal peptides into amines using the proteolytic enzymes and releasing amines such as methylamine, isobutylamine, putrescine, cadaverine, histamine, tyramine, and phenethylamine into the vaginal fluid increasing its pH (Cruden and Galask 1988; Sobel 2000; Wolrath *et al.* 2001). Amines in turn cause vaginal transudation and squamous epithelial cell exfoliation, leading to the vaginal discharge that is a typical clinical symptom of women with BV.

Even after decades of research, the aetiology of BV is still poorly understood. Major theories on the causes of BV revolve around *G. vaginalis*, lactobacilli, bacteriophages and the genetic predisposition of the host (Turovskiy *et al.* 2011). However, risks associated with BV are well documented. Many case control and cohort based studies have documented that BV increases the risk of preterm birth (Goldenberg *et al.* 2008; Hillier *et al.* 1995; Oleen-Burkey and Hillier 1995). BV is also implicated in causing pelvic inflammatory disease (PID). Anaerobic species associated with BV such as *Prevotella bivia,* other *Prevotella* species*,* and *Peptostreptococcus* species have been isolated from the fallopian tube tissue of women with acute PID (Soper *et al.* 1994). BV is also associated with post-abortion PID, post-hysterectomy cuff cellulitis, post-Caesarian endomyometritis as well as with cervical intraepithelial neoplasia (Sobel 2000).

Both symptomatic or "asymptomatic" BV (asymptomatic being defined as cases with high Nugent score, but with no clinical symptoms) are associated with both obstetric complications and gynecologic complications (Gibbs 2007). However, treatment of asymptomatic BV is recommended only in cases of women who are either at high risk for

preterm delivery or undergoing surgical procedures (2002).

The most common treatment options for BV include prescription medications such as oral metronidazole or clindamycin, intravaginal metronidazole gel or intravaginal clindamycin cream. There are other options available such vaginal hormone replacement treatment (Vagifem) and 'over-the-counter' treatments (Powell 2013). Other alternatives include triple-sulfa creams, erythromycin, tetracycline, acetic acid gel as well as many types of vaginal douches. Due to their poor efficacy, none of the above described alternatives have become popular in treating BV. Currently, the most effective treatment for BV is oral metronidazole therapy. Intravaginal metronidazole gel and intravaginal clindamycin cream have also been shown to be as effective as oral metronidazole (Ferris *et al.* 1995). However, about 30% of patients who initially responded to oral metronidazole therapy experience recurrence of symptoms within three months. The reason for recurrence is still poorly understood. Some of the explanations provided are the possibility of re-infection, presence of metronidazole resistant organisms (*Mycoplasma hominis*, *Mobiluncus curtisii*), the possible development of resistance by previously sensitive species (*G. vaginalis*) and protection of BV associated microbes through biofilms (Sobel 2000; Spiegel 1991). Effective treatment can be achieved for individuals by choosing optimal doses and varying treatment lengths, but tailored treatment approaches are required for some cohorts of women, which presents a significant clinical challenge (Koumans *et al.* 2002).

1.4 *Gardnerella vaginalis*

Gardnerella vaginalis is considered as one of the most important microorganisms as far as the genital tract health of women is concerned. It is sometimes referred to as the

hallmark organism of BV. This microbe has been studied for the past 65 years. First isolated in 1953 by Leopold (1953), initially named as *Haemophilus vaginalis* by Gardner and Dukes in 1955 (1955) and later as *Corynebacterium vaginale* by Zinnemann and Turner (1963), *G. vaginalis* was an organism classified with great difficulty. Due to its unusual cell wall, it required intense study employing methods such as DNA-DNA hybridization, and electron microscopy to establish that this organism was distinct from previously characterized Gram positive or Gram negative genera (Greenwood and Poickett 1980; Piot *et al.* 1980). The name '*Gardnerella*' was eventually proposed for this genus by Greenwood and Pickett (1980), supported by Piot *et al.* (1980). *G. vaginalis* remains the only species in the genus.

G. vaginalis (phylum Actinobacteria) is a fastidious, β-haemolytic, catalase-negative, non-capsulated, immobile pleomorphic rod. The Gram stain for *G. vaginalis* varies from Gram negative to Gram positive and hence is often described as 'Gram variable'. Ultrastructural examination of *G. vaginalis* ATCC 14018^T generated by Reyn *et al.* (1966) indicated well defined septa formed before cell division, which is indicative of Gram positive cells. Another characteristic feature of a Gram positive cell wall exhibited by *G. vaginalis* is the absence of diaminopimelic acid and lipopolysaccharide in the cell wall (Sadhu *et al.* 1989). Characteristics suggestive of a Gram negative cell wall of *G. vaginalis* are its amino acid composition, laminated structure and low peptidoglycan content (Catlin 1992; Piot *et al.* 1980).

G. vaginalis is routinely isolated from women with BV but is also present in women without BV. However, quantitative polymerase chain reaction (PCR) studies indicate high levels of *G. vaginalis* associated with BV diagnosis (Menard *et al.* 2008;

Zariffard *et al.* 2002). Other bacteria associated with BV such as genital mycoplasmas and strict anaerobes species such as BVAB1, BVAB2, BVAB3, *Atopobium*, *Leptotrichia*, *Megasphaera*, and *Prevotella* are isolated at varying rates from women with BV. Many conceptual model based theories exist on how *G. vaginalis* with the aid of its many virulence factors adheres to the vaginal epithelium of the host and successfully competes with the microbiota present in the vagina leading to pathogenesis of BV. *G. vaginalis* requires virulence factors for accomplishing three functions in the vagina: adherence to host receptor sites, production of cytotoxins specific for host cells, and biofilm formation. Some of the virulence factors produced by *G. vaginalis* include vaginolysin (Gelber *et al.* 2008), sialidase (Santiago *et al.* 2011b), bacteriocins (Teixeira *et al.* 2010), hemolysin (Cauci *et al.* 1993), and human lactoferrin binding protein (Jarosik and Land 2000). Vaginolysin is a human specific cholesterol-dependent cytolysin (CDC) that has the ability to form pores in target cell membranes resulting in lysis of target host cells (Gelber *et al.* 2008). Sialidase is a mucinase, specifically a neuraminidase that cleaves alpha-ketosidic linkages between glycosyl residues of glycoproteins, glycolipids and sialic acids. Hemolysin, a 59 kDa cytolysin characterized in *G. vaginalis,* is specific to human erythrocytes and is able to lyse the cells (Cauci *et al.* 1993). *G. vaginalis* is known to produce biofilm and biofilms produced by *G. vaginalis* are demonstrated to be present on the surface of the vaginal epithelium in women with BV (Patterson *et al.* 2010; Swidsinski *et al.* 2005). Biofilms produced by *G. vaginalis* have been implicated in sexual transmission of BV (Swidsinski *et al.* 2010). However, the protein factors involved in production of biofilms are still incompletely characterized.

As previously described, the Nugent score based on Gram stain is used to diagnose BV, and *G. vaginalis* abundance in the sample is one major factor determining the Nugent score. However, a high Nugent score does not always correspond to the presence of clinical symptoms. As more studies of vaginal microbiomes are conducted, *G. vaginalis* is becoming recognized as a heterogeneous group of organisms, and is also being accepted as a part of normal vaginal microbiota. Studies report that organisms grouped as *G. vaginalis* are phenotypically very diverse (Benito *et al.* 1986; Piot *et al.* 1984). *G. vaginalis* has been classified into eight biotypes based on the presence of *β*-galactosidase, lipase and hippurate hydrolysis ability (Piot *et al.* 1984), whereas Benito and associates have identified seventeen biotypes (Benito *et al.* 1986). Piot biotype 5 was found to be one of the most frequently isolated biotypes from healthy women (Aroutcheva *et al.* 2001). The same study reports Piot biotypes 7 and 8 as the most frequently isolated from BV patients with isolation rates of 32% (for biotype 7) and 20% (for biotype 8) (Aroutcheva *et al.* 2001). Interestingly, Piot biotype 5 has also been reported to be associated with BV (Benito *et al.* 1986). Some biotypes of *G. vaginalis* from both typing schemes have been associated with BV (Aroutcheva *et al.* 2001; Numanovic *et al.* 2008). Genotypic heterogeneity defined by different *Gardnerella*-like sequences detected in previous *cpn*60 sequence libraries (Hill *et al.* 2005a; Schellenberg *et al.* 2009a) of vaginal samples has also been described. Phenotypic and genotypic heterogeneity of the diverse group of organisms grouped under the binomial '*Gardnerella vaginalis*' needs to be better characterized and sub-categories must be resolved to permit the sub-classification of *G. vaginalis* and to understand if clinical significance can be associated with the sub-categories if resolved.

1.5 Mollicutes

Species of the class Mollicutes isolated from the urogenital tract of humans include *Mycoplasma fermentans, Mycoplasma genitalium, Mycoplasma hominis, Mycoplasma penetrans, Mycoplasma primatum, Mycoplasma spermatophilum, Ureaplasma parvum*, and *Ureaplasma urealyticum*. Mollicutes are well known to live as parasites on or in various animal and plant tissues. They are typically small in size (0.2 - 0.3 μm) and possess relatively very small genomes. For example, *Mycoplasma genitalium* has a genome size of 580 kb.

Genital mycoplasmas and ureaplasmas are strongly associated with BV, and with other female genital infections such as vaginitis, cervicitis, and PID. Review of the literature suggests that four Mollicutes species are significant predictors of genital tract health of women: *Mycoplasma hominis, Mycoplasma genitalium, Ureaplasma urealyticum* and *Ureaplasma parvum*. *Mycoplasma hominis* is an important organism in this group that is believed to be involved in BV, infertility, miscarriage, respiratory distress in newborns, and neonatal meningitis (Schlicht *et al.* 2004; Taylor-Robinson 1996). Ureaplasmas are associated with preterm labour, decreased sperm motility, intrauterine infection, lung disease, and neonatal pneumonia (Waites *et al.* 2005). These organisms are regularly detected in vaginal samples using culture-based methods or species-specific PCR.

Despite their apparent clinical significance and high prevalence in women of various health status, mycoplasmas and ureaplasmas are rarely detected in metagenomic studies, especially those using next-generation sequencing based methods. One possible explanation for the absence or under-representation of Mollicutes in metagenomic

libraries is that the primers used to amplify the target gene (usually 16S rRNA) are biased against amplification from certain taxa, including *Mycoplasma* and *Ureaplasma* (Hummelen *et al.* 2010). Mollicutes have not been detected in *cpn*60 libraries from the vaginal microbiome (Hill *et al.* 2005a; Schellenberg *et al.* 2009a). This is due to the absence of the *cpn*60 target gene, along with several other genes previously considered "essential", in virtually all *Mycoplasma* and *Ureaplasma*. The exception is *Mycoplasma genitalium*, which contains a *cpn*60 gene, but still has not been detected in studies. The capability of detecting and identifying Mollicutes in samples from the vaginal microbiome is essential to our goal of achieving comprehensive profiling methods for the evaluation of vaginal microbial environments. Methodologies to detect *Mycoplasma hominis, Mycoplasma genitalium, Ureaplasma parvum, Ureaplasma urealyticum* and any other significant Mollicutes in vaginal samples need to be developed and validated by adopting and modifying methods already established for their detection in tissue culture and the identification of cultured isolates.

1.6 Preterm premature rupture of membranes

Preterm birth is one of the foremost perinatal challenges worldwide that often results in serious implications for the mother and the baby. Preterm birth overburdens medical facilities and systems in respective countries. The preterm birth rate in North America is about 10.6% and the annual cost of care for preterm infants in the United States alone is greater than $26 billion (Christopher *et al.* 2009).

In pregnancy, disturbances of the vaginal microbial community are associated with the release of inflammatory cytokines and chemokines (Cauci *et al.* 2002). Microbial endotoxins and cytokines produced by the host stimulate the production of

prostaglandins and matrix-degrading enzymes (Roberton *et al.* 2005), which can lead to the stimulation of uterine contraction or weakening of the amniotic membrane and eventually its rupture (Mascagni and Miller 2004). Preterm rupture of membrane (PROM) is defined as rupture of membrane/amniotic sac before the beginning of labour. When PROM occurs before 37 weeks of pregnancy, it is called preterm premature rupture of membrane (PPROM), which can lead to preterm birth. Women having PPROM can be categorized into three groups based on the outcome of time of delivery following membrane rupture: women who deliver immediately following membrane rupture (< 48 hrs), women with extended latencies between 48 hrs and ≤ 7 days following membrane rupture, and women that deliver > 7 days following membrane rupture. However, the factors that determine the duration of latency are not well understood.

Among the major pathophysiological pathways contributing to PPROM and the outcomes of premature birth are abnormal vaginal microbiota and associated maternal infection and inflammatory response (Green *et al.* 2005). Though there is limited understanding of the pathophysiological pathways contributing to the etiology of preterm birth, intrauterine infection resulting from an abnormal/unhealthy vaginal microbial community is thought to play a significant role in triggering PPROM, especially at lower gestational ages.

Microbes such as *Gardnerella, Mycoplasma, Ureaplasma, Mobiluncus* and other aerotolerant and anaerobic organisms are associated with BV and other genital tract issues. Intrauterine infection due to microbes present in vagina is considered an important risk factor for premature preterm rupture of membrane (PPROM). Many virulence factors from BV-associated organisms have been characterized (Wiggins *et al.* 2001) including

mucinases, cytolysins, prolidases, siderophores, human-lactoferrin binding proteins, lipases, and phospholipases.

Mucinases are hydrolyzing enzymes secreted by microorganisms to lyse and degrade vaginal mucus. Mucinases have been recognized for a long time, but only some like sialidases have been studied extensively. Other known mucinases in the human vagina include glycosidases, proteases, and sulfatases. Mucinases are produced mostly by anaerobic bacteria and probably aid in their successful colonization during BV by enabling adherence to vaginal epithelial cells. Wiggins *et al.* (2001) suggested that glycosidase enzymes result in cervical mucin degradation, enabling colonization of the upper reproductive tract and chorioamnionitis in pregnant women. Sialidases are relatively well-characterized enzymes with mucinase and neuraminidase activity that have been described to affect the collagen synthesis ability of fibroblasts (Aalto *et al.* 1974). Sialidase activity of vaginal fluid is highly correlated with BV, and is used as a marker for diagnosing BV (Briselden *et al.* 1992; Cauci *et al.* 2005; Shujatullah *et al.* 2010). Collagen content of the fetal membrane is directly responsible for the tensile strength of the membrane. Decreased collagen content, altered collagen structure and increased collagenolytic activity are proposed to be associated with PPROM (Epstein *et al.* 1998). Prolidases and proteases are also linked to collagen metabolism and prolidases have been proposed to predict preterm birth in combination with pH and sialidase (Cauci *et al.* 2005).

Cytolysins, pore-forming exotoxins that lyse animal cells, are important virulence factors produced by many Gram negative bacteria (Welch 1991). Vaginolysin is one such human-specific cholesterol dependent cytolysin, characterized in *G. vaginalis* (Gelber *et*

al. 2008). Prolidases are capable of modulating the activity of immune factors and high levels of prolidase were inversely correlated with the host's ability to elicit an appropriate immune response (Schoonmaker *et al.* 1991; Vanhoof *et al.* 1995). One of the ways that intrauterine infection can cause preterm membrane rupture is by eliciting an immune response that increases the production of prostaglandins, in turn causing membrane rupture.

Phospholipases are enzymes that can degrade phospholipids present on the cell surface. Phospholipase A_2 is reported to be able to release arachidonic acid from membrane phospholipids of the amnion. Arachidonic acids are precursors to prostaglandins. During an immune response, cytokines in the uterus produce cyclooxygenase II, which can convert arachidonic acid to prostaglandins. Prostaglandins induce uterine contractions and stimulation, which can lead to membrane rupture. Phospholiapase A_2 activity is reported to be higher in pregnant women with BV (Jones and Al-Mushrif 1997) .

Despite a growing understanding of specific microbial virulence factors and the pathophysiology of PPROM, there is still little understanding of what kinds of vaginal microbial communities indicate a risk of PPROM or predict outcomes following PPROM. This is due to a number of challenges faced in this area of research. Major challenges include overly simplistic diagnostic tools for detecting abnormal vaginal microbiota, and failure to discriminate BV from other conditions such as aerobic vaginitis. Research is also significantly impeded by difficulty in obtaining samples from women with PPROM due to the unpredictability of PPROM and the complex clinical aspects of this potentially devastating event. Another important challenge is the

uniqueness of vaginal microbiomes. The Human Microbiome Project claims that every person has a unique microbial community on the skin (Fierer *et al.* 2010). Researchers in the vaginal microbiome area report difficulty in identifying "core" organisms associated with the microbiome (Ravel *et al.* 2011), suggesting that every woman may have a unique vaginal microbiome with varying richness and diversity of taxa. A high-resolution picture of the vaginal microbiome and the ability to resolve its members to the species level, that in turn can be linked to differences in their genotypic and phenotypic properties is needed to decipher any signature pattern that might help clinicians predict PPROM and its outcomes.

OBJECTIVES

The overall objective of this research project is to obtain a better understanding and higher resolution view of the vaginal microbial community that can be applied to defining the relationship of this community to health or disease. Specific objectives that follow are:

- Determine if *G. vaginalis cpn*60 sequence diversity observed in metagenomic studies of vaginal microbiota represents phenotypically distinct subpopulations.
- Comprehensively define *G. vaginalis* subgroups through comparative genomics, proteomics, detection of putative virulence factors and identification of biomarkers.
- Investigate the association of vaginal microbiome structure with maternal and neonatal outcomes following preterm premature rupture of membrane (PPROM).

CHAPTER 2 - Resolution and characterization of distinct *cpn*60-based subgroups of *Gardnerella vaginalis* in the vaginal microbiota

Copyright

Citation

Paramel Jayaprakash T, Schellenberg JJ, Hill JE (2012) Resolution and Characterization of Distinct *cpn*60-Based Subgroups of *Gardnerella vaginalis* in the Vaginal Microbiota. PLoS ONE 7(8): e43009. doi:10.1371/journal.pone.0043009

Author Contributions

Conceived and designed the experiments: TPJ JJS JEH. Performed the experiments: TPJ. Analyzed the data: TPJ JJS JEH. Wrote the paper: TPJ JJS JEH.

Second Chapter Transition

In this manuscript I addressed the first objective of the project, which was to resolve genotypic and phenotypic heterogeneity of *G. vaginalis*, a hallmark group of organisms associated with BV. I investigated whether biotypes defined by the earliest and most popular biotyping scheme (Piot biotyping (Piot *et al.* 1984)) are consistent with genotypes described for the genus. The ultimate motivation behind resolving the heterogeneity is to understand any clinical significance that may be associated with resolved groups. Hence, distribution of the resolved subgroups was investigated in a cohort of Kenyan women grouped into three based on their Nugent score: normal, intermediate and BV.

2.1 Abstract

Bacterial vaginosis (BV), characterized by a shift of the vaginal microbiota from a *Lactobacillus* dominated community to a dense biofilm containing a complex mixture of organisms, is an important risk factor in poor reproductive health outcomes. The Nugent score, based on Gram stain, is used to diagnose BV and *G. vaginalis* abundance in the sample is one factor determining Nugent score. A high Nugent score is indicative of BV but does not always correspond to the presence of clinical symptoms. *G. vaginalis* is recognized as a heterogeneous group of organisms, which can also be part of the normal, healthy vaginal microbiome. In addition, asymptomatic BV and non-*Gardnerella* types of BV are being recognized. In an attempt to resolve the heterogeneous group of *G. vaginalis*, a phylogenetic tree of *cpn*60 universal target sequences from *G. vaginalis* isolates was constructed that indicates the existence of four subgroups of *G. vaginalis*. This subdivision, supported by whole genome similarity calculation of representative strains using JSpecies, demonstrates that these subgroups may represent different species. The *cpn*60 subgroupings did not correspond with the Piot biotyping scheme, but did show consistency with ARDRA genotyping and sialidase gene presence. Isolates from all four subgroups produced biofilm *in vitro*. We also investigated the distribution of *G. vaginalis* subgroups in vaginal samples from Kenyan women with Nugent scores consistent with BV, Intermediate and Normal microbiota (n=44). All subgroups of *G. vaginalis* were detected in these women, with a significant difference ($z = -3.372$, $n = 39$, $p = 0.001$) in frequency of *G. vaginalis* subgroup B between BV and Normal groups. Establishment of a quantifiable relationship between *G. vaginalis* subgroup distribution and clinical status could have significant diagnostic implications.

2.2 Introduction

Gardnerella vaginalis, first isolated by Leopold in 1953 (Leopold 1953), has long been recognized in vaginal samples and has been identified by several names, including *Haemophilus vaginalis* by Gardner and Dukes in 1955 (Gardner and Dukes 1955). Further characterization based on metabolic requirements and Gram staining led to its reclassification as *Corynebacterium vaginale* (Zinnemann and Turner 1963). The proposal to create the genus *Gardnerella* and allocation of *Corynebacterium vaginale* and *Haemophilus vaginalis* to this new taxon as *Gardnerella vaginalis* was put forward by Greenwood and Pickett (Greenwood and Poickett 1980), based on a taxonomic study that utilized DNA-DNA hybridization, biochemical analysis of the cell wall, and electron microscopy.

G. vaginalis is strongly associated with BV, and is one of the most frequently isolated bacteria from women with symptoms of BV (Demba *et al.* 2005a, b; Hillier *et al.* 1991; Schellenberg *et al.* 2011b). Abundance of *G. vaginalis* in vaginal samples has also been associated with infertility and preterm labour (Menard *et al.* 2010). *G. vaginalis* has also been isolated from urine and blood and is associated with bacteremia, osteomyelitis and cervical cancer (Graham *et al.* 2009; Johnson and Boustouller 1987; Mikamo *et al.* 1999). However, recent studies of vaginal microbiota indicate that *G. vaginalis* can also be a part of the vaginal microbiota in clinically healthy women (Hill *et al.* 2005a; Schellenberg *et al.* 2011b; Tosun *et al.* 2007).

G. vaginalis is recognized as a diverse taxon, both phenotypically and genotypically (Benito *et al.* 1986; Piot *et al.* 1984; Tosun *et al.* 2007). Eight biotypes of *G. vaginalis* have been identified by Piot *et al*. based on the presence of β-galactosidase,

lipase and hippurate hydrolysis activities (Piot *et al.* 1984), whereas Benito *et al*. identified seventeen biotypes based on these characteristics in addition to fermentation of xylose, arabinose and galactose (Benito *et al.* 1986). Phenotypic diversity within *G. vaginalis* has also been described in terms of virulence factors, particularly production of sialidase (Santiago *et al.* 2011b) and formation of biofilms (Swidsinski *et al.* 2010). Genetic heterogeneity within *G. vaginalis* has been demonstrated using amplified ribosomal DNA restriction analysis (ARDRA) (Ingianni *et al.* 1997). Santiago *et al*. (Santiago *et al.* 2011b), identified three ARDRA genotypes of *G. vaginalis*, of which only two genotypes (genotypes 1 and 3) produced sialidase. However, like biotyping schemes, ARDRA can only be performed on isolates.

Genotype diversity is apparent in whole genome studies and in metagenomic studies of the human vaginal microbiome based on 16S rRNA (Hummelen *et al.* 2010) or *cpn*60 (Hill *et al.* 2005a; Schellenberg *et al.* 2011b). Hummelen *et al.* (2010) reported the presence of four types of *G. vaginalis* sequences differing by a single nucleotide within the 16S rRNA V6 region. Four clusters of *G. vaginalis* sequences, ranging between 89 and 100% sequence identity to the type strain (ATCC 14018^{T}), were observed in a *cpn*60-based study of clinically healthy women by Hill *et al.* (2005a) and followed up in a larger study by Schellenberg *et al*. (2011b).

Previous work by our research group has demonstrated that *cpn*60 universal target sequences can resolve phenotypically distinct strains or ecotypes within an intestinal microbial community, and that these sequences are also excellent predictors of whole genome sequence relationships (Verbeke *et al.* 2011; Vermette *et al.* 2010). The gene encoding the universal 60 kDa chaperonin protein (*cpn*60) is an established target for

detection and identification of microorganisms, as well as gene-based metagenomic studies of complex microbial communities, including the vaginal microbiome (Chaban *et al.* 2012; Desai *et al.* 2012; Desai *et al.* 2009; Dumonceaux *et al.* 2006a; Dumonceaux *et al.* 2006b; Fernando *et al.* 2010; Hill *et al.* 2005a; Hill *et al.* 2005b; Hill *et al.* 2002; Schellenberg *et al.* 2009a; Schellenberg *et al.* 2011a; Schellenberg *et al.* 2011b). An approximately 555 bp region corresponding to nucleotides 274-828 of the *E. coli cpn*60 gene can be amplified with degenerate, universal PCR primers (Goh *et al.* 1996; Hill *et al.* 2006b). This universal target (UT) region is phylogenetically informative, providing more discriminating power than 16S rRNA to differentiate organisms, even at the sub-species or strain level (Blaiotta *et al.* 2008; Brousseau *et al.* 2001; Goh *et al.* 2000; Hill *et al.* 2006a; Hung *et al.* 2010; Jian *et al.* 2001; Minana-Galbis *et al.* 2009; Sakamoto and Ohkuma 2010; Sakamoto *et al.* 2010; Vermette *et al.* 2010). A highly curated reference database of chaperonin sequences, cpnDB, supports *cpn*60 based applications (Hill *et al.* 2004).

Given the observed phenotypic diversity (especially virulence factors), genotypic diversity, and the presence of *G. vaginalis* in women regardless of clinical status, it is critical to improve our understanding of the clinical significance of these different strains. In order to accomplish this most effectively and to lay the foundation for the development of more informative diagnostic tools for women's health, direct culture-independent analysis of vaginal samples, exploiting a genetic target that facilitates robust resolution is required.

The objective of the current study was to investigate if previously observed *cpn*60 based subgroups of *G. vaginalis* are consistent with other (phenotypic) classification

systems and/or available whole genome sequences, and to investigate the distribution of *cpn*60 defined subgroups of *G. vaginalis* in women with and without BV. Our results demonstrate that the *cpn*60 universal target sequence differentiates distinct subgroups within *G. vaginalis* and that only one of these subgroups (Subgroup B: Piot biotype 5, sialidase positive and ARDRA genotype 1) was found to be significantly more abundant in women with BV (high Nugent score) than women with normal vaginal microbiota in a retrospective analysis of metagenomic profiles of Kenyan women.

2.3 Materials and Methods

Bacterial isolates

G. vaginalis ATCC 14018 (type strain) and ATCC 49145 were obtained from the American Type Culture Collection (Manassas, VA). Eight additional strains were isolated from Kenyan (N170, N165, N160, N158, N153, N148, N144, N143, N137, N134, N101, and N72) or Canadian (W11) women as described previously (Schellenberg *et al.* 2011b). *G. vaginalis* isolates were cultured using the following media: ATCC #1685 broth (with or without 1% (w/v) glucose), Brain Heart Infusion broth (BHI) with 1% (w/v) glucose, egg yolk agar (Piot *et al.* 1984) and Columbia agar with 5% sheep blood (BD, Mississauga, ON). The GasPak EZ Pouch System (BD, Mississauga, ON) was used to provide anaerobic conditions.

DNA extraction and PCR

DNA was extracted from isolates using a phenol-chloroform extraction method and was stored at -20 °C. Primers used in the study are listed in Table 2.1. *cpn*60 UT PCR amplicons were produced for direct sequencing, using universal primers H729 and

H730 as described previously (Hill *et al.* 2006b). Primers JH0315 and JH0316 were designed based on the 16S rRNA sequence from *G. vaginalis* ATCC 14018. Amplification with these primers was carried out by incubating the reactions at 94 °C for 3 minutes, followed by 40 cycles of 94 °C for 30 sec, 52 °C for 1 min and 72 °C for 90 sec, and completed with a final extension of 10 min at 72 °C. Sialidase gene presence was assessed by amplifying the sialidase gene using primers GVSI forward and GVSI reverse (Santiago *et al.* 2011b). Vaginolysin gene sequences were amplified using primers V1 and V2 as previously described (Gelber *et al.* 2008).

Table 2.1. Primers used in the study.

Primer name	Sequence (5'-3')	Reference
H729	CGC CAG GGT TTT CCC AGT CAC GAC GAI III GCI GGI GAY GGI ACI ACI AC	(Hill *et al.* 2006b)
H730	AGC GGA TAA CAA TTT CAC ACA GGA YKI YKI TCI CCR AAI CCI GGI GCY TT	(Hill *et al.* 2006b)
JH0315	ATT CTG GCT CAG GAT GAA	This study
JH0316	GCT ACC TTG TTA CGA CTT AG	This study
GVSI forward	GAC GAC GGC GAA TGG CAC GA	(Santiago *et al.* 2011b)
GVSI reverse	AGT CGC ACT CCG CGC AAG TC	(Santiago *et al.* 2011b)
V1	ATG CAG CGA AGC ATG CCA TGC	(Gelber *et al.* 2008)
V2	TCA GTC GTT CTT TAC AGT TTC	(Gelber *et al.* 2008)
GV10F	GGT TCG ATT CTG GCT CAG	(Santiago *et al.* 2011b)
ωMB	TAC CTT GTT ACG ACT TCG TCC CA	(Santiago *et al.* 2011b)

Phenotyping of *G. vaginalis* isolates

Representative *G. vaginalis* isolates with unique *cpn*60 sequences were phenotyped using the Piot typing scheme using assays for hippurate hydrolysis, β-galactosidase and lipase activity as described previously (Piot *et al.* 1984). *Lactobacillus crispatus* was used as a negative control for the lipase assay. *G. vaginalis* ATCC 14018^T was used as a positive control for all biochemical assays.

Biofilm formation

Isolates of *G. vaginalis* were cultured from -80 °C stocks for 72 hrs, and subcultured for 48 hrs on Columbia sheep blood agar plates anaerobically at 37 °C. A loopful of culture for each isolate was used to inoculate 4 ml of either ATCC broth #1685 supplemented with 1% (w/v) glucose or Brain Heart Infusion supplemented with 1% (w/v) glucose (BHIG) and incubated anaerobically for 48 hrs at 37 °C. Broth cultures were diluted 1:100 in media and 200 µl of diluted culture was added to individual wells of a 96-well tissue culture plate and incubated anaerobically for 48 hrs at 37 °C. Qualitative assessment of biofilm formation was done by washing off the planktonic cells and staining the wells with 1% crystal violet solution to visualize any biofilm.

ARDRA genotyping

G. vaginalis isolates were genotyped using amplified rDNA restriction analysis (ARDRA) (Ingianni *et al.* 1997; Santiago *et al.* 2011b). Full-length 16S rRNA gene sequences were amplified using primers GV10F and ωMB (Table 2.1). PCR products were purified (QIAquick PCR Purification Kit, Qiagen, Inc., Toronto, ON) and subjected

to overnight restriction digestion using TaqI (Life Technologies, Inc., Burlington, ON). The digestion products were resolved on a 1.5% agarose gel at 140 volts for 2 hrs. *In silico* ARDRA was performed for some strains for which only published genome sequence information was available (not the isolates themselves) by extracting full length 16S rRNA gene sequences from published whole genome sequences and then restricting the sequence using the program remap within the EMBOSS software suite (Rice *et al.* 2000).

Sequence sources

Published genome sequences of *G. vaginalis* used in the study, either completed or in progress, were downloaded from NCBI's Genome database. The metagenomic *cpn*60 sequences and *cpn*60 UT sequences of *G. vaginalis* isolates used were from a previously published study of vaginal microbiota of commercial sex workers in Kenya (Schellenberg *et al.* 2011b).

Phylogenetic analysis

*cpn*60 UT sequences obtained from whole genome sequences of *G. vaginalis* reference strains or amplified from cultured clinical isolates (Schellenberg *et al.* 2011a) were used to construct a phylogenetic tree, using *Alloscardovia omnicolens* CCUG 34444 as a root. Sequences were aligned using ClustalW (gap opening penalty = 10, gap extension penalty = 0.10) (Thompson *et al.* 1994), followed by utilization of the Phylip software package (Felsenstein 1993) to calculate a distance matrix using dnadist and construct a tree using neighbor. The final tree was obtained from the bootstrapped consensus of 100 trees and was visualized using Dendroscope (Huson *et al.* 2007).

Statistical analysis

All statistical analyses were done using SPSS Statistics, version 19.0. For the analysis of assignment of assembled reads of Nairobi metagenomic data set to *G. vaginalis* subgroups, one-way ANOVA was done, followed by Post-hoc analysis by Tukey's test. Statistical analysis for the distribution of *G. vaginalis* subgroups in Nairobi women was done by Kruskal-Wallis H test, followed by Mann-Whitney U test.

2.4 Results

*cpn*60-based resolution of *G. vaginalis* subgroups

A phylogenetic tree of *G. vaginalis cpn*60 UT sequences is shown in Figure 2.1A. This tree was used to select landmark sequences for further comparisons and for subgrouping of *G. vaginalis* metagenomic sequences. The discreteness of the subgroups was further supported by demonstration of a bimodal distribution of pairwise distances (inter- and intra-subgroup) between strains (Figure 2.1B). The four subgroups resolved were designated A, B, C and D and two representative sequences from each subgroup were selected as representatives to capture the maximum phylogenetic distance represented by the phylogenetic tree.

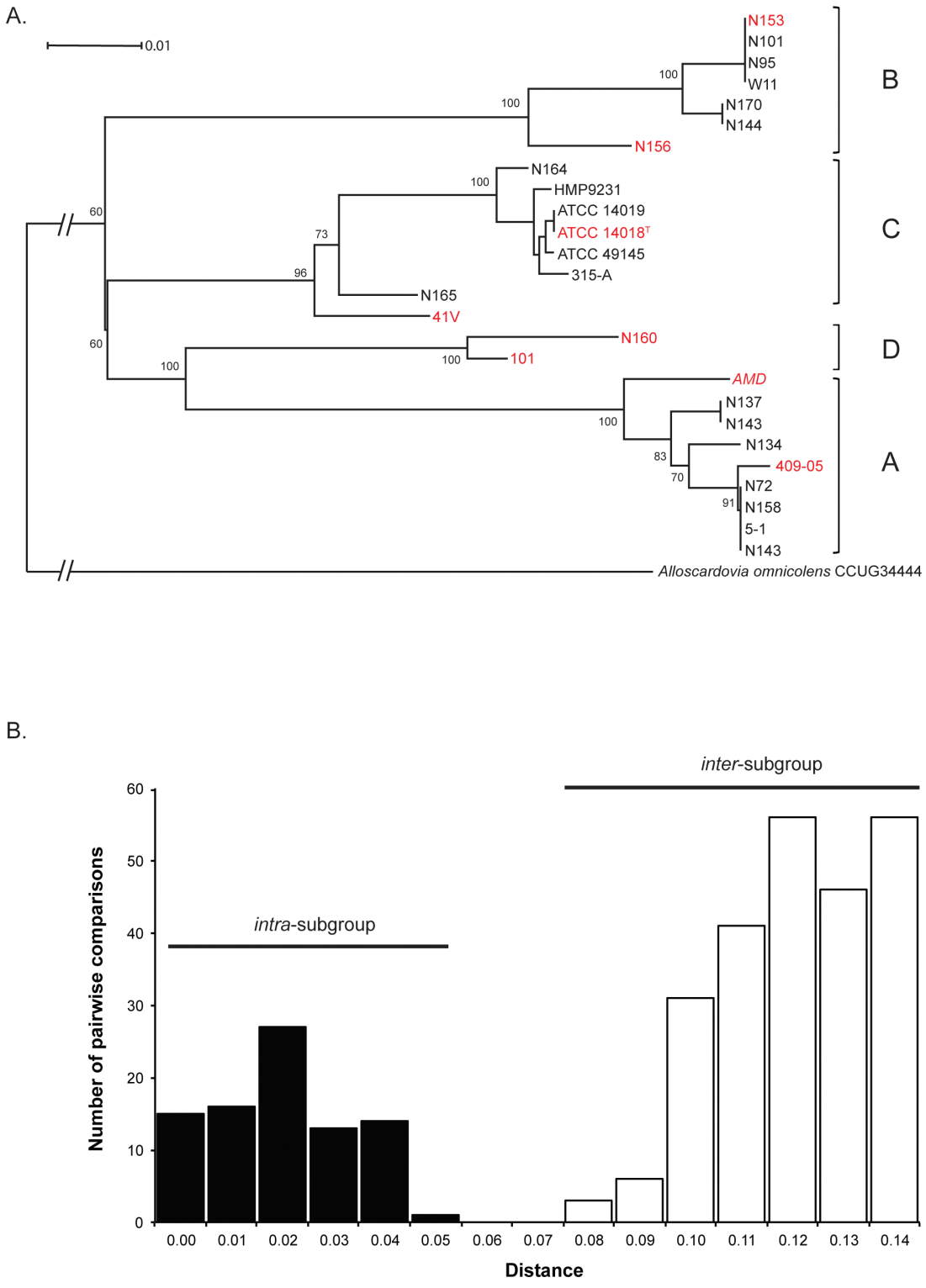


Figure 2.1. *cpn*60 UT sequence-based subgroups of *G. vaginalis*.
A. Phylogenetic tree of *G. vaginalis*-like *cpn*60 UT sequences comprising four distinct clades: A, B, C and D. Bootstrap values for each node are indicated. 101, 315-A, 41V, 409-05, 5-1, AMD, ATCC 14018T, ATCC 14019, and HMP9231 are *G. vaginalis* isolates with whole genome sequence information available in Genbank (Accession numbers AEJD00000000, AFDI00000000, AEJE00000000, CP001849, ADAN00000000, ADAM00000000, ADNB00000000, CP002104 and CP002725 respectively). Isolates

with names starting with "N" are isolates from Kenyan women from Schellenberg *et al.* (2011a). W11 was isolated from a Canadian woman (Schellenberg, Unpublished). Sequences highlighted in red were used as representatives of the subgroups in the distribution analysis of metagenomic sequence data. B. Pairwise distances for the 26 *G. vaginalis cpn*60 UT sequences included in the phylogenetic analysis. Distances for both inter-subgroup comparisons (white bars) and intra-subgroup comparisons (black bars) are indicated.

Whole genome, *cpn*60 and 16S rRNA comparisons

Whole genome similarity calculations of representative strains of *G. vaginalis* for which there was either complete or partial genome sequence data available were calculated with JSpecies (Table 2.2). Within subgroups, pairwise Average Nucleotide Identity by MUMmer (ANIm) values were >95% and *cpn*60 identities were ≥96%, while between subgroups ANIm values were <90% and *cpn*60 identities were ≤92%. Pairwise identities based on16S rRNA were all 98-100%. Pairwise *cpn*60 and 16S rRNA sequence identities for representative isolates with unique *cpn*60 sequences are shown in Table 2.3. The pairwise sequence identity for *cpn*60 gene sequences and 16S rRNA gene sequences for isolates within *G. vaginalis* subgroups are >96% and >97%, respectively. Between *G. vaginalis* subgroups, pairwise identities ranged between 87% and 93% for *cpn*60 sequences and 97% and 100% for 16S rRNA sequences.

Table 2.2. ANIm (first row), *cpn*60 UT sequence identity (second row) and 16S rRNA sequence identity (third row) for representatives of *G. vaginalis* subgroups A, C and D, for which whole genome sequence data was available. No whole genome sequence is available for a subgroup B strain.

		A		C					D
		AMD	5-1	ATCC 14019	41V	ATCC 14018	315-A	HMP9321	101
A	409-05	95.86	98.2	89.09	88.97	89.13	88.85	89.1	88.56
		97	99	88	90	88	88	88	91
		99	100	98	98	98	98	98	99
	AMD		95.7	89.06	88.64	89	88.98	89.59	88.51
			97	89	90	89	88	89	91
			99	98	98	98	98	98	99
	5-1			89.31	88.89	89.29	88.96	89.96	88.91
				88	91	88	88	88	91
				98	98	98	98	98	99
C	ATCC 14019				95.91	99.79	98.19	98.35	88.29
					96	100	99	99	92
					99	100	100	100	98
	41V					95.87	96.04	96.07	88.66
						96	96	96	92
						99	99	99	98
	ATCC 14018						98.13	98.27	88.22
							99	99	92
							100	100	98
	315-A							98.38	88.45
								99	92
								100	98
	HMP9321								89.19
									92
									98

Table 2.3. Pairwise sequence identity of *cpn*60 (first row of each column) and full-length 16S rRNA (second row) for *G. vaginalis* isolates. ND = Not done. Only representative study isolates with unique *cpn*60 sequences are included.

	A					B			C								D	
	N158	N134	AMD	5-1	N137	N156	N153	N144	ATCC 14019	N165	N164	41V	ATCC14018	315-A	HMP9231	ATCC49145	N160	101
409-05	99	98	97	99	98	90	87	87	88	90	89	90	88	88	88	88	90	91
	99	99	99	100	99	ND	99	99	98	98	ND	98	98	98	98	98	99	99
N158		98	97	100	98	90	87	87	88	90	89	91	88	88	88	88	90	91
		99	99	99	99	ND	99	99	98	98	ND	98	98	98	98	98	99	99
N134			97	98	98	89	87	88	88	90	89	90	88	88	88	88	90	91
			99	99	99	ND	99	99	98	98	ND	98	98	98	98	98	99	99
AMD				97	97	89	87	88	89	91	89	90	89	88	89	89	90	91
				99	100	ND	99	99	98	98	ND	98	98	98	98	98	99	99
5-1					98	90	87	87	88	90	89	91	88	88	88	88	90	91
					99	ND	99	99	98	98	ND	98	98	98	98	98	99	99
N137						89	87	88	89	91	90	91	89	89	89	89	90	91
						ND	99	99	98	98	ND	98	98	98	98	98	99	99
N156							96	96	88	90	89	89	88	89	89	88	88	88
							ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
N153								98	89	90	89	89	89	90	89	89	88	88
								99	99	99	ND	99	99	99	99	99	99	99
N144									89	91	90	90	89	90	90	90	88	88
									99	99	ND	99	99	99	99	99	99	98
ATCC14019										97	99	96	100	99	99	99	90	92
										99	ND	99	100	100	100	99	98	98
N165											97	97	97	96	96	96	91	93
											ND	99	99	99	99	99	98	98
N164												96	99	98	99	98	91	92
												ND	ND	ND	ND	ND	ND	ND
41V													96	96	96	96	91	92
													99	99	99	99	98	98
ATCC14018^T														99	99	99	90	92
														100	100	99	98	98
315-A															99	99	90	92
															100	99	98	98
HMP9231																99	91	92
																99	98	98
ATCC49145																	90	92
																	98	98
N160																		98
																		99

Biotyping, sialidase, ARDRA genotyping and biofilm production

Results for the biotyping assays and genotyping are shown in Table 2.4. Piot biotypes 1, 2, 5, 7 and 8 were identified among the isolate collection, but no consistent pattern of biotype distribution and *cpn*60 subgroup was observed. All subgroup C isolates were lipase positive. Subgroup B and C isolates were sialidase positive and ARDRA genotype 1, whereas subgroup A isolates were sialidase negative and genotype 2. Subgroup D isolates differed in sialidase (N160 was negative, strain 101 positive) and ARDRA genotype (N160 was genotype 2, strain 101 was predicted to be genotype 1 based on *in silico* restriction analysis). All isolates produced biofilm in BHIG by 48 hrs, and substantial variability in biofilm production was observed in the two media tested (BHIG and ATCC broth #1685 with 1% glucose) (Figure 2.2). Both subgroup B isolates (N144 and N153) formed biofilm in both media, although the biofilm formed in BHIG was more extensive, completely coating the well. In subgroups A and C, at least one of the isolates failed to produce any visible biofilm in ATCC broth #1685.

Table 2.4. Piot biotype, sialidase and ARDRA characterization of *G. vaginalis* isolates (representatives of study isolates with unique *cpn*60 sequences) and published whole genome sequences.

*cpn*60 subgroup	Isolate[2]	Piot Biotype[1]				Sialidase	ARDRA[3]
		L	B	H	Biotype		
C	41V	ND	ND	ND	ND	+	1
	N165	+	+	-	8	+	1
	ATCC 14018	+	+	+	1	+	1
	ATCC 49145	+	+	-	8	+	1
	ATCC 14019	ND	ND	ND	ND	+	1
B	N144	-	-	+	5	+	1
	N153	-	-	+	5	+	1
D	101	ND	ND	ND	ND	+	1
	N160	-	-	+	5	-	2
A	AMD	ND	ND	ND	ND	-	2
	N137	-	-	-	7	-	2
	N134	-	-	+	5	-	2
	409-05	ND	ND	ND	ND	-	2
	5-1	ND	ND	ND	ND	-	2
	N158	+	-	+	2	-	2

[1]L = Lipase, B = β-galactosidase, H = Hippurate hydrolase, ND = not done.
[2]Study isolates N156 (subgroup B) and N164 (subgroup C) were not included in the biotyping analysis since they were not reliably cultured as pure isolates.
[3]In cases where no isolates were available to us for culture, ARDRA genotypes for some strains (41V, ATCC 14019, 101, AMD, 409-05, and 5-1) were obtained by *in silico* analysis as described in the text.

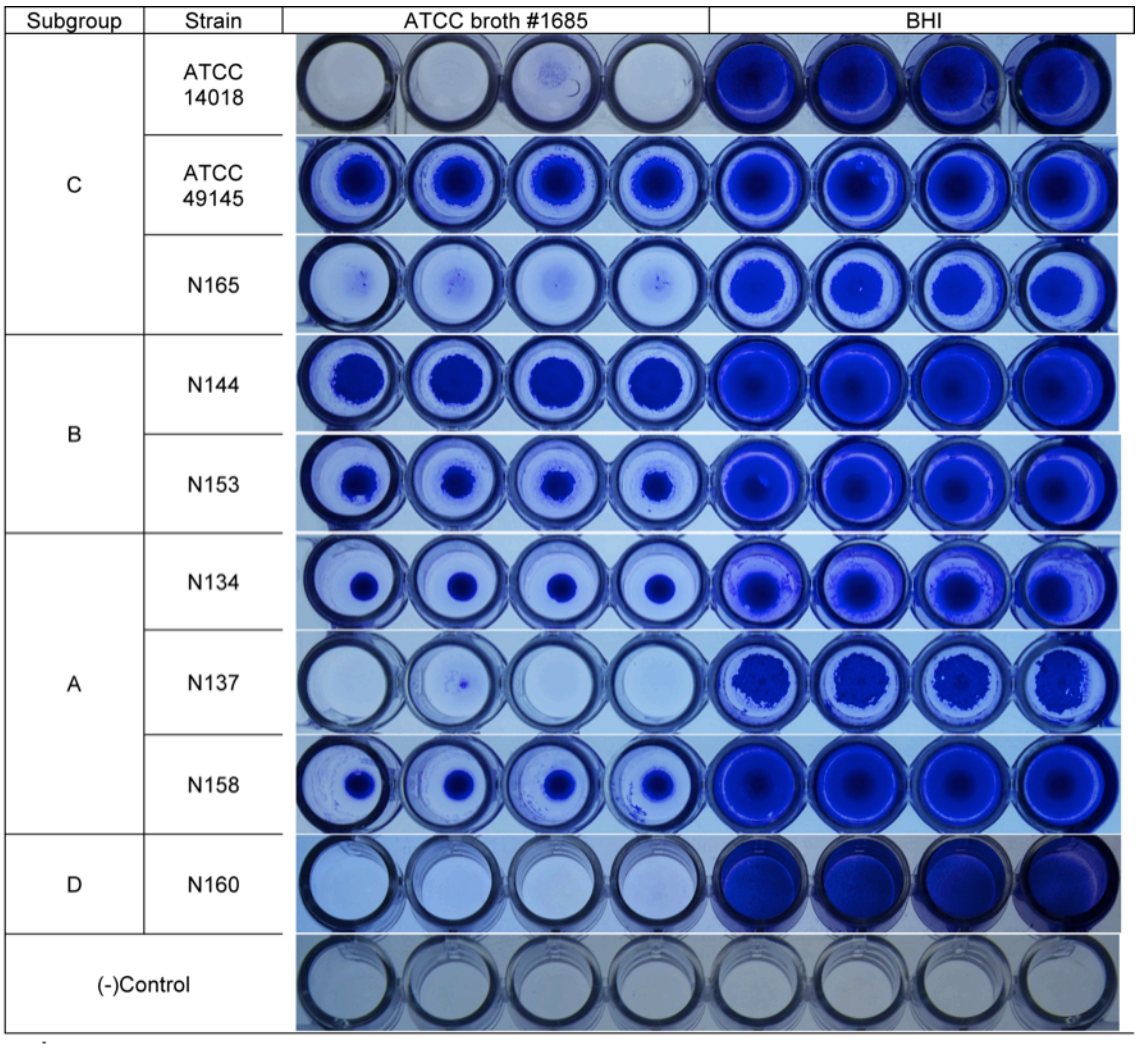


Figure 2.2. Biofilm formation by *G. vaginalis.*
Isolates were cultured in 96-well plate in two different media: ATCC broth #1685 and BHI, stained at 48 hrs, after removal of planktonic cells.

We also attempted to detect vaginolysin gene sequences in the isolates selected for phenotypic analysis. A PCR product of the expected size of 1,551 bp was obtained for only three isolates (ATCC 14018^T, ATCC 49145 and N153). An amplicon of 1,200 bp was amplified from three others (N134, N137 and N158), but sequence analysis indicated a mixture of products, suggesting that this product was the result of non-target sequence amplification. Isolates N165 and N144 did not yield any product after repeated attempts. The vaginolysin sequence from ATCC 49145 was identical to ATCC 14018 and only 89% identical to N153.

Distribution of *G. vaginalis* subgroups in Kenyan women

A previously published *cpn*60 metagenomic dataset was used to investigate distribution of *cpn*60-based *G. vaginalis* subgroups in vaginal microbiome profiles derived from samples classified as BV, Intermediate or Normal based on Nugent score (Schellenberg *et al.* 2011a). All unique sequences assembled from the study data (n=831 OTU) were compared using watered-BLAST (Schellenberg *et al.* 2009a) to a reference database of *cpn*60 sequences containing one representative of each species in cpnDB (cpnDB_nr; www.cpndb.ca) and two representatives of each *G. vaginalis* subgroup as indicated in Figure 1. All assembled sequences with any of the *G. vaginalis* reference sequences as their best match, and meeting the minimum requirement for identification as a *cpn*60 sequence (60% identity over ≥100 nucleotides) were included in the analysis of distribution (n=93). For 84/93 of the assembled metagenomic sequences, the top two hits were to the same *G. vaginalis* subgroup. Identities for each query and its top two hits (medians for 93 queries were 96.6% and 92.9% respectively) were significantly ($p <$

0.0001) higher than identity to the third (median 89.4%) through eighth best hits (Figure 2.3). Thirty-one queries had sequence identities <95% to their best match.

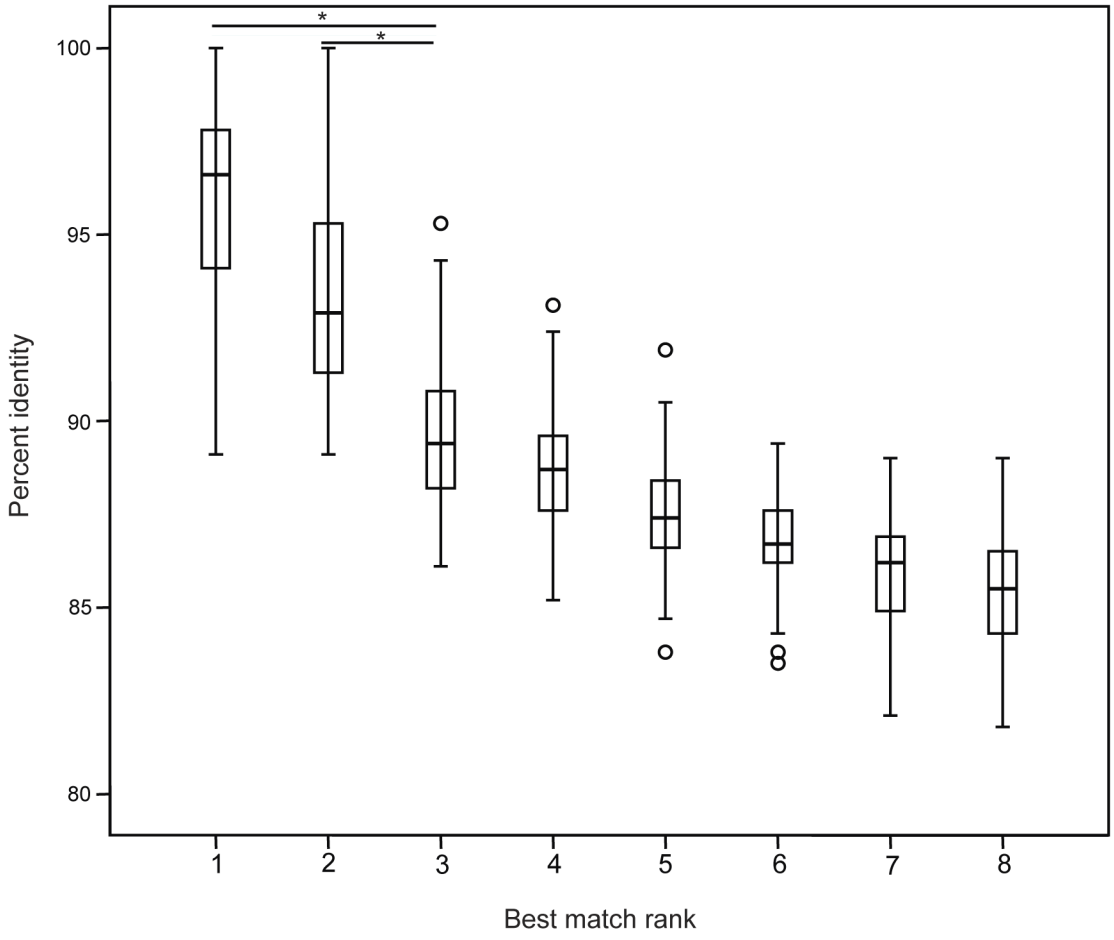


Figure 2.3. Percent identity of metagenomic sequences to *G. vaginalis* reference strains.
Distribution of percent identity of *G. vaginalis* metagenomic sequences to their first through eighth best matches (x axis values 1 to 8) among representative sequences of the *G. vaginalis* subgroups. The reference database included two representatives of each subgroup, indicated in Figure 1. Significant differences in percent identity ($p<0.0001$) are indicated by *.

Once metagenomic sequences were assigned to a subgroup (based on the best watered-BLAST match) the distribution of the subgroups among the vaginal microbiomes of women diagnosed as BV (n = 20), Intermediate (n = 5) and Normal (n = 19), based on Nugent score was determined (Figure 2.4). The sequence read frequencies used in this analysis were normalized to the median library size of 15,000 reads (Schellenberg *et al.* 2011a). All vaginal microbiota libraries contained sequences corresponding to more than one subgroup of *G. vaginalis*. Out of 44 libraries sequenced, 41, 43, 43 and 27 contained *G. vaginalis* subgroup A, B, C and D respectively. The majority of libraries (25/44) contained sequence from all four subgroups. The next most prevalent combination was A+B+C (n=14 libraries), while other combinations were present in the remaining four samples (n=1 for A+C, n=2 for B+C, n=1 for B+C+D, n=1 for A+B+D). The difference in frequencies of *G. vaginalis* subgroup sequence reads was tested using Kruskal-Wallis H test (SPSS Statistics, version 17.0). A significant difference (χ^2 value = 12.329, df = 2, p = 0.002) was observed only for subgroup B sequences. Pairwise comparison on subgroup B sequences between the clinical groups (BV, Intermediate, and Normal) was analyzed by Mann-Whitney U test and results showed a significantly greater abundance of subgroup B sequences in BV compared to Normal (z = -3.372, n = 39, p = 0.001).

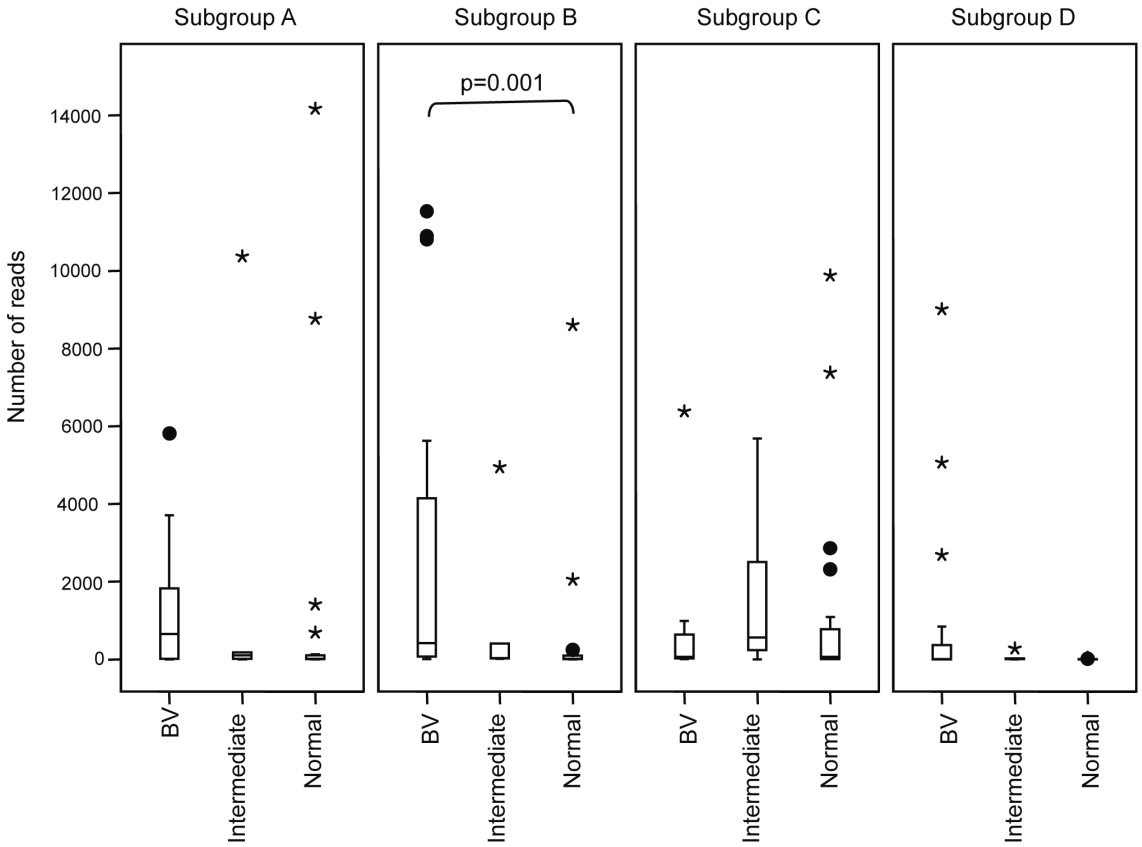


Figure 2.4. Distribution of *G. vaginalis* subgroups in African women.
Relative abundance of sequence reads corresponding to *Gardnerella* subgroups (scaled to median library size of 15,000 reads) among clinical categories (BV, Intermediate and Normal, based on Nugent score). Boxplots were created for each *Gardnerella* subtype and p values calculated based on non-parametric significance tests (Mann-Whitney U test) using SPSS Statistics version 19.0.

2.5 Discussion

Bacterial vaginosis is the most commonly reported vaginal infection (Fethers *et al.* 2008; Ness *et al.* 2004). BV can be diagnosed clinically using Amsel's criteria, which include presence of homogenous vaginal discharge, a vaginal pH of greater than 4.5, positive whiff test (production of a fishy odour on addition of 10% KOH to vaginal sample), and also presence of clue cells in at least 20% of the total cell count (Amsel *et al.* 1983). The Nugent score is another commonly used diagnostic tool for BV. To calculate the Nugent score, a Gram stained vaginal smear is assessed for the relative abundance of various bacterial morphotypes including Gram-positive large rods, Gram-negative/Gram-variable rods and curved Gram-variable rods. With increasing numbers of survey studies in which Nugent scores of clinically normal women are determined, the phenomenon of "asymptomatic BV" has been widely observed (Gibbs 2007; Leitich and Kiss 2007; Mastrobattista *et al.* 2000). These women have high Nugent scores, but do not have symptoms of BV. The clinical significance of asymptomatic BV is unknown. Since the presence of *G. vaginalis* is one of the key determinants of Nugent score, one possible explanation for asymptomatic BV is the presence of large numbers of non-pathogenic *G. vaginalis* or other species with similar Gram stain morphology. If this is true, then the detection of *G. vaginalis* in general may be of questionable diagnostic value. Resolution of this important issue requires the investigation of distribution of *G. vaginalis* lineages among women in a variety of clinical cohorts. Tackling this on a large scale requires culture-independent tools that provide resolution of phenotypically distinct *G. vaginalis* subgroups when applied directly to clinical samples.

Sequence diversity within *G. vaginalis* has been reported in metagenomic studies of the vaginal microbiome based on 16S rRNA and *cpn*60 gene targets, and four subdivisions have been reported in several recent studies (Hill *et al.* 2005a; Hummelen *et al.* 2010; Schellenberg *et al.* 2011b). The subdivision observed by Hummelen *et al.* (2010) was based on single nucleotide differences within the V6 region of 16S rRNA. The much lower *cpn*60 sequence identity between these subgroups (≤93% versus ≥98% identity for 16S rRNA) facilitated the identification of vaginal isolates corresponding to these subgroups, demonstrating that the metagenomic studies had revealed real biological diversity and not artifactual diversity resulting from PCR, sequencing and or data assembly (Figure 2.1). Overestimation of microbial diversity in metagenomic sequencing studies is an ongoing concern (Reeder and Knight 2009). So the identification and characterization of actual isolates corresponding to metagenomic sequences is reassuring and further supports the value of the *cpn*60 universal target for resolution of diversity at species- and strain-level (Vermette *et al.* 2010).

Whole genome DNA–DNA hybridization persists as the gold standard method for defining bacterial species (1988) although it remains unpopular due to its technical demands (Cho and Tiedje 2001). DNA sequence data is increasingly relied upon to support species definition and resolution, and recently whole genome sequence comparison has been suggested as a new gold standard (Richter and Rossello-Mora 2009). Richter & Rosselló-Mora (2009) demonstrated that average nucleotide identity (ANI) values correlate well with DNA-DNA hybridization results and suggest that an ANI values greater than ≈95-96%, calculated by either BLAST or the MUMmer rapid aligning tool, were indicative of bacteria of the same species. Another alternative was

proposed by Ziegler, who developed a computational algorithm based on sequence of three genes (*recN, rpoA* and *thdF*) that corresponds well to the conclusions of DNA-DNA hybridization data (Zeigler 2003). In that study, the 16S rRNA gene, widely used for identifying bacterial species and metagenomic studies, was found to have the lowest correlation between sequence identity and genome sequence identity. Subsequently, Verbeke *et al.* (2011) demonstrated that a *cpn*60 UT sequence alone could predict whole genome identity as well as the three-gene model. The ease of amplifying and sequencing the *cpn*60 UT from bacteria, the curated reference database of chaperonin sequences (cpnDB, www.cpndb.ca) (Hill *et al.* 2004), and the ability of the *cpn*60 UT to predict whole genome sequence similarity make it the ideal target for studies of *Gardnerella*, or any other bacterial taxon for which subspecies resolution is of interest.

Our results show a strong relationship between *cpn*60 UT sequence identity and whole genome comparison with the ANIm algorithm in JSpecies (Table 2.2). In fact, our results suggest that subgroups A, C and D of *G. vaginalis* meet the whole genome sequence-based criteria for designation as different species. Although no whole genome sequence is available for a subgroup B isolate, the *cpn*60 sequence data for isolates in this group certainly support a similar species level status for this group (Table 2.3). Complete genomes of nine *G. vaginalis* strains were determined at the time of writing, and the fact that none of them belongs to subgroup B is interesting. Our experience with culturing of subgroup B isolates suggest that their conspicuous absence from the genome sequence database is most likely due to the fact that members of this subgroup, unlike the others, only grow in anaerobic conditions and do not grow in 7% CO_2, which is the atmosphere recommended for routine isolation of *G. vaginalis* (Holt 1994).

The heterogeneity of the *G. vaginalis* taxon is well documented based on application of biotyping schemes. Some of the biotypes of *G. vaginalis* from both the Piot and Benito biotyping schemes have been associated with BV (Aroutcheva *et al.* 2001; Numanovic *et al.* 2008). Piot biotypes 1, 4 and 5 are the most frequently isolated regardless of BV status (Tosun *et al.* 2007) and biotype 5 has been reported to be predominantly associated with healthy vaginal ecosystems (Aroutcheva *et al.* 2001). Piot biotypes 7 and 8 have been reported as the most frequently isolated from BV patients with isolation rate of 32% and 20%, respectively (Aroutcheva *et al.* 2001). Of the seven isolates characterized in this study, four were Piot biotype 5, supporting previous observations of the prevalence of this biotype (Table 2.4). Otherwise, biotyping results were not consistent with sialidase, ARDRA or observations of association with BV in the Kenyan cohort. Sialidase activity is recognized as a virulence factor in *G. vaginalis* and is the basis for a chromogenic, BV Blue Kit for BV diagnosis (Kampan *et al.* 2011). Furthermore, although we did not provide evidence of sialidase enzymatic activity in our isolates, sialidase gene presence has previously been correlated with the ability of an isolate to produce active sialidase enzyme (Santiago *et al.* 2011b). We observed a consistent relationship between sialidase production and ARDRA genotype in that all genotype 1 isolates were sialidase positive and all genotype 2 isolates were sialidase negative, in agreement with previous observations (Santiago *et al.* 2011b).

Vaginolysin is a protein toxin belonging to the cholesterol-dependent cytolysin family of toxins that has been previously identified in *G. vaginalis* (Gelber *et al.* 2008). To detect this purported virulence factor in the study isolates, we employed previously published PCR primers designed based on the type strain, ATCC 14018. The primers

failed to amplify the target sequence from most study isolates, and among the isolates for which we did generate sequence (ATCC 14018, ATCC 49145 and N153), we observed only 89% nucleotide sequence identity between some isolates (N153 vs. either the type strain or ATCC 49145). These results suggest that these primers may be too specific for general application in *G. vaginalis,* rather than indicating the absence of a vaginolysin gene in the other strains included in the study.

Subgroups of *G. vaginalis* were not evenly distributed among vaginal microbiomes diagnosed as BV, intermediate or normal based on Nugent score (Figure 2.4). Although *G. vaginalis* sequences were ubiquitous in the study group, and most women hosted multiple subgroups of *G. vaginalis*, only sialidase positive subgroup B was significantly more abundant in BV than normal samples. Analysis of pH and clue cells in these samples showed, as expected, a negative correlation of pH and Nugent score and a positive correlation of clue cells and Nugent score (data not show). An obvious and immediate question is whether subgroup B or any other subgroup is differentially associated with symptomatic and asymptomatic BV. Although we were unable to stratify our current data by symptoms (discharge, odour), relatively low *cpn*60 sequence identities facilitate robust differentiation of *G. vaginalis* subgroups (Figure 2.1 and 2.3) making it an ideal target to exploit in culture-independent approaches to addressing these questions in future studies. High throughput sequencing of *cpn*60 amplicons (Schellenberg *et al.* 2009a), bead-based hybridization assays (Dumonceaux *et al.* 2009) and quantitative real-time PCR methods (Chaban *et al.* 2009) have all been developed based on *cpn*60 UT sequences and offer powerful tools for investigation of microbial diversity at, and below, the species level.

The results of our work support previous observations of genotypic and phenotypic diversity within *G. vaginalis* and we have been successful in using *cpn*60 UT sequences for robust classification of available *G. vaginalis* strains into four subgroups. We have also provided evidence that supports the eventual reclassification of subgroups as different species of *Gardnerella*. The degree of *cpn*60 UT and whole genome sequence diversity within this taxon is beyond that associated with "ecotypes" (Cohan 2002) or strains and suggests that reclassification may be warranted. However, additional genotypic and phenotypic analysis of additional isolates will be required to make this case. The cpn60 UT sequence offers a robust tool for identification of subgroups within *G. vaginalis* that may not be discernable using other targets. This feature of the *cpn*60 target will facilitate future efforts to expand diagnostic panels for rapid, high throughput characterization and improved resolution of species and strain distribution in the vaginal microbiome.

CHAPTER 3 - Identification of putative virulence factors in *cpn*60-defined subgroups of *Gardnerella vaginalis* using phenotypic and molecular profiling

Citation

Jayaprakash TP, Schellenberg JJ, Lessard C, Hill JE. Identification of putative virulence factors in *cpn*60-defined subgroups of *Gardnerella vaginalis* using phenotypic and molecular profiling.

Author Contributions

Conceived and designed the experiments: TPJ JJS CL JEH. Performed the experiments: TPJ JJS. Analyzed the data: TPJ JJS JEH. Wrote the paper: TPJ JJS CL JEH.

Third Chapter Transition

In the previous chapter, four subgroups of *G. vaginalis* based on *cpn*60 UT were defined. During the same time another group (Ahmed *et al.* 2012) reported that *G. vaginalis* can be subdivided into four clades based on the whole genome sequences of 17 isolates within this genus. In this chapter, I have reconciled *cpn*60 subgroups of *G. vaginalis* with the Ahmed *et al.* clades (Ahmed *et al.* 2012). The other objective I accomplished in this chapter is to resolve the four subgroups based on their proteome profiles. These results indicated that subgroups are recognizable using more than one method of resolving bacterial species. The next important information required about these subgroups is the identification of phenotypic and molecular factors that might indicate their clinical significance. Hence, sialidase enzymatic activity of subgroups and also their antimicrobial susceptibility to metronidazole and clindamycin were investigated. The final objective of this chapter was to identify any subgroup specific protein biomarker candidates with diagnostic potential.

3.1 Abstract

Gardnerella vaginalis is a ubiquitous member of the vaginal microbiome that plays an ambiguous role in the pathogenesis of bacterial vaginosis (BV). Although phenotypic and phylogenetic diversity in this organism is well known, recent work has confirmed four genotypically distinct subgroups (A-D), which are unevenly distributed among women with BV as determined by Nugent score. To further characterize the subgroups, we assessed several phenotypic and molecular factors of *G. vaginalis* subgroups, using representative isolates of each subgroup grown to stationary phase. Presence and activity of sialidase was determined through gene-specific PCR and the filter-spot enzymatic assay, antimicrobial resistance patterns for metronidazole and clindamycin were determined using E-strips, and whole-cell proteomic profiles were generated by SDS-PAGE. Sialidase gene sequences were detected in all subgroups, however enzymatic activity was detected only in subgroup B. Proteomic profiles of isolates within each subgroup formed unambiguous clusters, with subgroup B isolates expressing apparently large amounts of a 43.7 kDa protein identified using mass spectrometry as a putative ABC transporter substrate binding protein (COG2182: maltose-binding periplasmic proteins/domains). A clear definition of *G. vaginalis* subgroups in terms of molecular and phenotypic traits, as well as their distribution in the context of the vaginal microbiome, will be critical to a better understanding of the pathogenesis of BV. Characterization of specific protein markers for each subgroup is likely to lead to the development of novel diagnostic methods for this common but poorly understood clinical entity.

3.2 Introduction

The most common vaginal condition affecting reproductive aged women is bacterial vaginosis (BV), characterized by a shift in the composition of the vaginal microbiota from a *Lactobacillus* dominated community to a polymicrobial community, dominated by anaerobes (Spiegel 1991). *G. vaginalis* is a hallmark organism of BV, however it is also detected routinely in clinically healthy women (Harwich *et al.* 2010). The phenotypic heterogeneity of *G. vaginalis* is well known, and several biotyping and genotyping schemes have been developed to describe this diversity (Benito *et al.* 1986; Piot *et al.* 1984; Santiago *et al.* 2011b; Tosun *et al.* 2007). More recently, it has been demonstrated that this taxon consists of four genotypically and phenotypically distinct subgroups that may constitute separate species. These subgroups can be discriminated on the basis of concatenated sequences of 332 genes common to *G. vaginalis* isolates (Ahmed *et al.* 2012), or by sequencing a single 552 bp region of *cpn*60, also known as the "universal target" or UT (Paramel Jayaprakash *et al.* 2012). It has not yet been determined if these genotypic subgrouping schemes are consistent with each other or if clinically significant phenotypic characteristics are differentially distributed among subgroups. There is only limited information on the association between any particular subgroup or clade of *G. vaginalis* and BV status, including the results of our own previous work (Paramel Jayaprakash *et al.* 2012). Recently, quantitative PCR using subgroup-specific probes has shown associations between *G. vaginalis* subgroups and Nugent score in 60 women attending a vaginitis clinic in the United States (Balashov *et al.* 2014).

Sialidase activity is an important virulence factor associated with BV, aerobic vaginitis, and adverse pregnancy outcomes (Cauci and Culhane 2011; Marconi *et al.* 2013; Zhang *et al.* 2002). Although this activity is widely detected in *G. vaginalis,* the trait is not common to all isolates, and expression levels are variable among sialidase positive isolates (Lewis *et al.* 2013). Some recent studies have investigated the presence of sialidase genes in relation to sialidase activity and molecular subgroups (Pleckaityte *et al.* 2012a; Santiago *et al.* 2011b). Our previous work has demonstrated that sialidase genes could be detected in *cpn*60 defined subgroups B and C, however enzymatic activity was not determined (Paramel Jayaprakash *et al.* 2012). We are unaware of any previous studies that have addressed antibiotic resistance or whole-cell protein profiles of *G. vaginalis* subgroups.

The objectives of the current study were to reconcile current genomic and phenotypic descriptions of *G. vaginalis* subgroups defined by *cpn*60 and whole genome sequences, and to expand on the phenotypic descriptions of *G. vaginalis* subgroups with the goal of identifying characteristics that can be exploited in further investigations of their clinical significance, and their detection and identification. Our results demonstrate that clades defined by Ahmed *et al.* (2012), based on supragenome genome analyses of 17 *G. vaginalis* genomes, correspond to *cpn*60 subgroups, and furthermore that *cpn*60 subgroups are distinguishable based on whole cell protein profiles. The detection of sialidase enzymatic activity in only one of the *cpn*60-defined subgroups and the identification of subgroup-specific proteins that likely play a role in nutrient uptake support the hypothesis that *G. vaginalis* subgroups are also ecologically distinct, having

different relationships with the host and the other constituents of the vaginal microbial community.

3.3 Materials and methods

Bacterial cultures

All *G. vaginalis* isolates included in this study are listed in Table 3.1. *G. vaginalis* isolates ATCC 49145 and ATCC 14018^T were obtained from the American Type Culture Collection. All other isolates were from vaginal samples of Kenyan and Canadian women, as previously described (Schellenberg *et al.* 2012). The current study focuses on a set of 13 isolates, including 9 that have been previously characterized (Paramel Jayaprakash *et al.* 2012). These isolates represent four *cpn*60 defined subgroups of *G. vaginalis*: subgroup A (isolates N137, N134, N158), subgroup B (N144, N170, N95, W11, N101, N153), subgroup C (N165, ATCC 14018^T, ATCC 49145) and subgroup D (N160). A new set of 35 isolates, identified to the subgroup level by *cpn*60 UT sequencing as described previously by Hill *et al.* (Hill *et al.* 2006a) was also examined for confirmatory tests. Freezer stocks in skim milk-glucose or ATCC broth #1685 (NYC III medium) with 10% glycerol (v/v) were revived on Columbia blood agar or Mueller-Hinton blood agar (MHB), both with 5% sheep blood (BD, Mississauga, ON) and incubated anaerobically at 37°C using the GasPak EZ Pouch System (BD, Mississauga, ON). Isolated colonies were sub-cultured in ATCC broth #1685 and incubated as above. Stationary phase was determined by measuring absorbance of the culture at 600 nm and comparison to previously determined growth curves.

Table 3.1: *G. vaginalis* isolates included in the study. NA represents 'Not available'.

*cpn*60 subgroup	Ahmed *et al.* clade	Isolate	Isolate available	cpnDB ID	Whole genome accession number (Bioproject)
A	4	N134	Yes	b15982	NA
		N137	Yes	b15981	NA
		N143	Yes	b26701	NA
		N148	Yes	b26702	NA
		N158	Yes	b26703	NA
		N72	Yes	b15953	NA
		409-05	No	b15920	PRJNA31001
		AMD	No	b16964	PRJNA40893
		5_1	No	b17423	PRJNA40895
		B476	No	b21788	PRJNA42441
B	2	N101	Yes	b26700	PRJNA265097
		N144	Yes	b15978	NA
		N156	Yes	b15975	NA
		N170	Yes	b26704	NA
		N95	Yes	b26706	PRJNA265092
		W11	Yes	b26707	PRJNA265103
		JCP7659	No	NA	PRJNA181314
		JCP7719	No	NA	PRJNA181316
		JCP8017A	No	NA	PRJNA181317
		JCP8017B	No	NA	PRJNA181318
		JCP8066	No	NA	PRJNA181319
		JCP8070	No	NA	PRJNA181320
		JCP8151A	No	NA	PRJNA181322
		JCP8151B	No	NA	PRJNA181323
		JCP8522	No	NA	PRJNA181326
		N153	Yes	b15977	PRJNA265102
		B513	No	b21784	PRJNA42449
		B482	No	b21783	PRJNA42451
C	1	ATCC 49145	Yes	b26699	NA
		N164	Yes	b15970	NA
		N165	Yes	b15969	NA
		JCP7275	No	NA	PRJNA181312
		JCP7276	No	NA	PRJNA181313
		JCP7672	No	NA	PRJNA181315
		JCP8108	No	NA	PRJNA181321
		41V	No	b18714	PRJNA200477
		ATCC 14019	No	b13658	PRJNA31473
		B472	No	b21791	PRJNA42431
		B473	No	b21790	PRJNA42437
		B474	No	b21789	PRJNA42437
		B477	No	b21787	PRJNA42443
		B478	No	b21786	PRJNA42445
		ATCC 14018^{T}	Yes	b291	PRJNA46675
		HMP9321	No	b19071	PRJNA51067
		315-A	No	b19054	PRJNA52049
D	3	N160	Yes	b15973	NA
		B479	No	b21785	PRJNA42447
		B483	No	b21782	PRJNA42453
		B512	No	b21781	PRJNA42455
		101	No	b18713	PRJNA53359
OUTLIER	OUTLIER	JCP8481A	No	NA	PRJNA181324
OUTLIER	OUTLIER	JCP8481B	No	NA	PRJNA181325

Phylogenetic analysis and whole genome sequence comparisons

Phylogenetic analysis was performed by aligning sequences using ClustalW (gap opening penalty = 10, gap extension penalty = 0.10) (Thompson *et al.* 1994). Bootstrap values were determined using 100 iterations of the alignment, and distance matrices calculated using the F84 algorithm of *dnadist* in the PHYLIP software package (Felsenstein 1996). Phylogenetic trees were constructed using *neighbor* and *consense* in the PHYLIP package, and visualized using Dendroscope (Huson *et al.* 2007). For whole genome sequence comparisons, pairwise Average Nucleotide Identity (ANI) values between whole genome sequences were calculated using the MUMmer algorithm (ANIm) within the JSpecies software package for intergenome similarity comparisons (Richter and Rossello-Mora 2009).

Sialidase detection, Piot biotyping and antimicrobial susceptibility testing

PCR for the detection of the sialidase gene was performed using primers GVSI forward and GVSI reverse as previously described (Santiago *et al.* 2011b). A modified filter paper spot test was used for detecting sialidase enzyme activity of *G. vaginalis* isolates (Moncla and Braham 1989) with the fluorogenic substrate 2'-(4-methylumbelliferyl)-α–D-N-acetylneuraminic acid sodium salt hydrate (Sigma-Aldrich Canada, Oakville, ON). Substrate was dissolved in water to a concentration of 0.015% w/v (1 mg in 6.6 mL) and 180 μL aliquots stored at -20°C. Prior to the assay, thawed aliquots were diluted with 20 μL 1M sodium acetate (pH 5.8) and 10 μL added to filter spots. Isolates were sub-cultured on Columbia blood agar as described above and harvested with a sterile swab into 2 mL autoclaved saline (0.85% NaCl, pH 7.4) until

turbidity was equivalent to 5 McFarland. Sialidase activity was determined by visualizing filter spots under UV light after addition of 20 μL diluted culture and incubation for 15 min in the dark at 37°C. Based on preliminary experiments, ATCC 14018^T and isolate W11 were used as negative and positive controls respectively. To ensure that enzymatic activity was responsible for the positive reaction, the assay was repeated with diluted cultures that had been boiled for 30 min (enzyme-negative control). Biotyping was carried out on selected isolates using the Piot scheme (Piot *et al.* 1984), including assays for hippurate hydrolysis, *β*-galactosidase and lipase activity as previously described (Piot *et al.* 1984). Susceptibility testing for clindamycin and metronidazole was performed using E-test strips (bioMérieux, St. Laurent, Canada) as recommended by the manufacturer, with cultures diluted to 0.5 McFarland units and spread for confluent growth on MHB plates at 37°C for 48 hrs.

Protein preparation and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Protein extracts for 13 representative isolates were obtained from three replicate cultures. Bacterial cells were harvested from stationary phase cultures by centrifugation at 3900 × g for 5 min, washed three times with PBS and re-suspended in SDS lysis buffer, containing Tris (0.5 M, pH 6.8), SDS (8.8% w/v), glycerol (35.6% v/v), and bromophenol blue (0.3% w/v). Cell suspensions were sonicated (6 pulses of 30 sec, with each round followed by 2 min of cooling on ice) and cell debris was removed by centrifugation at 20,000 × g for 30 min at 4°C. Supernatants containing extracted proteins were stored at -80°C until analysis. Total protein concentration was determined using the

DC Protein Assay Kit II (Bio-Rad, Mississauga, ON). Protein extracts (75 µg total protein per isolate) were resolved on a 12% SDS-PAGE gel containing three lanes of size markers (Precision Plus Protein Standards, Bio-Rad, Mississauga, ON), stained with Coomassie brilliant blue overnight and destained for 20 min. Gels were scanned and images saved as .tiff files for analysis of band patterns using gel comparison software (GelCompar II V6.1, Applied Maths Inc. Austin, TX). Dendrograms were generated using UPGMA clustering with the Dice co-efficient, and 0.6% optimization and 0.8% position tolerance values. For more precise estimation of protein sizes and band intensities, selected extracts were resolved on an Agilent Bioanalyzer 2100 using the Protein 80 chip according to the manufacturer's instructions.

Protein identification by mass spectrometry

Bands corresponding to proteins apparently differentially expressed among *G. vaginalis* subgroups were excised. Mass spectrometry was performed for peptide mass determination using a Quadrupole Time-Of-Flight (Q-TOF) Global Ultima mass spectrometer (Micromass, Manchester, UK) at the National Research Council of Canada, Saskatoon, SK. RAW mass spectra data files obtained for each sample were processed with ProteinLynx Global Server 2.4 (PLGS 2.4, Waters) to generate PKL files containing peptide mass/charge ratio and intensity values. Each PKL data file was submitted to Mascot (Matrix Science Ltd., London, UK) for comparison of peptides to the SwissProt protein database, NCBI protein database and either to the predicted proteomes of *G. vaginalis* isolates N153 (*cpn*60 subgroup B, NCBI BioProject PRJNA265102) or B479 (1500E)(*cpn*60 subgroup D, GCA_000263595.1). The predicted proteome for each

genome was generated using *getorf* within the EMBOSS package (Rice *et al.* 2000) to identify all predicted open reading frames of at least 100 nucleotides.

3.4 Results

Groupings of *G. vaginalis* by *cpn*60 universal target and whole genome sequences

In order to reconcile the genome sequence-based grouping scheme described by Ahmed *et al.* (2012) with the *cpn*60 universal target based scheme described by Jayaprakash *et al.* (2012), phylogenetic analysis of *cpn*60 universal target sequences (552 bp) extracted from whole genome sequences described in Ahmed *et al.* (2012), *cpn*60 universal target sequences from Jayaprakash *et al.* (2012) and also from isolates described in a study by Lewis *et al*. (2013) was performed (Figure 3.1). Results of the analysis indicate that the four "clades" identified based on 332 concatenated genes are consistent with *cpn*60 subgroups, with *cpn*60 subgroup A corresponding to clade 4, subgroup B to clade 2, subgroup C to clade 1 and subgroup D to Clade 3 (Figure 3.1). Isolates JCP8481A and JCP8481B (Human Microbiome Project), which were not included in either previous study, shared a node (99% bootstrap support) with subgroup A, although they were only 89 - 91% identical to other subgroup A isolates. Pairwise comparisons of *cpn*60 UT sequences of *G. vaginalis* isolates within and between subgroups results in a bimodal distribution of inter- and intra-subgroup similarity values (Figure 3.2A). The only overlap between distributions occurs at 94% identity, a value representing 7 pairwise comparisons (6 intra-group and 1 inter-group). Pairwise Average Nucleotide Identity values calculated using the MUMmer algorithm exhibited a similar, bimodal distribution (Figure 3.2B). Within each subgroup, ANIm values were ≥ 95%, while between subgroup values ranged between 88% and 91% identity.

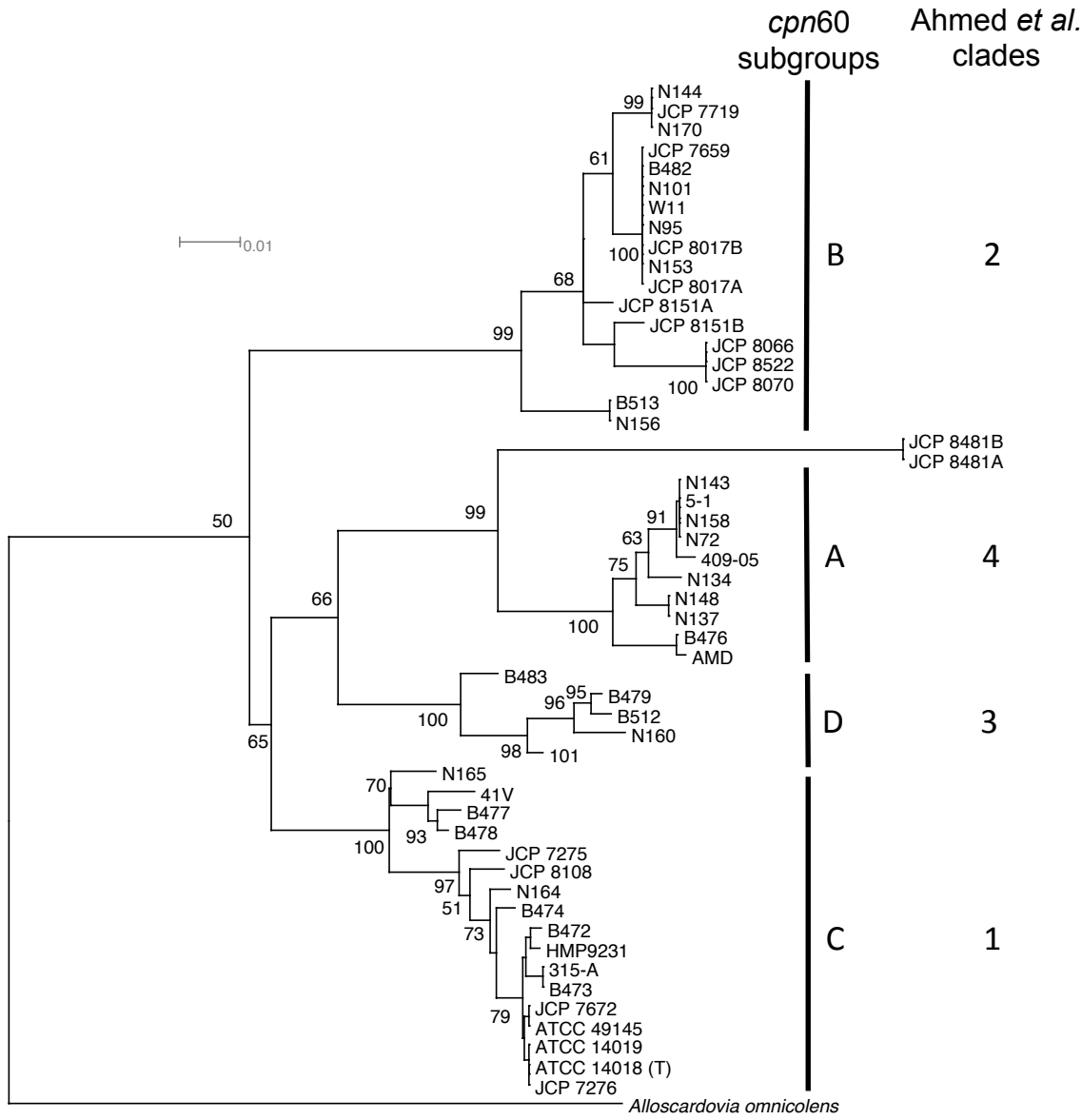


Figure 3.1. Phylogenetic analysis of the *cpn*60 UT sequences of *Gardnerella vaginalis.*
Phylogenetic tree of nucleotide sequences (alignment length = 552 bp) of *cpn*60 sequences from *G. vaginalis* (n=53) from four subgroups (A – D). The tree is rooted with *Alloscardovia omnicolens*. Sequences from isolates with names starting with B were obtained from Ahmed *et al.* (2012). All other sequences were either obtained from the JGI IMG/ER database or from NCBI Genbank.

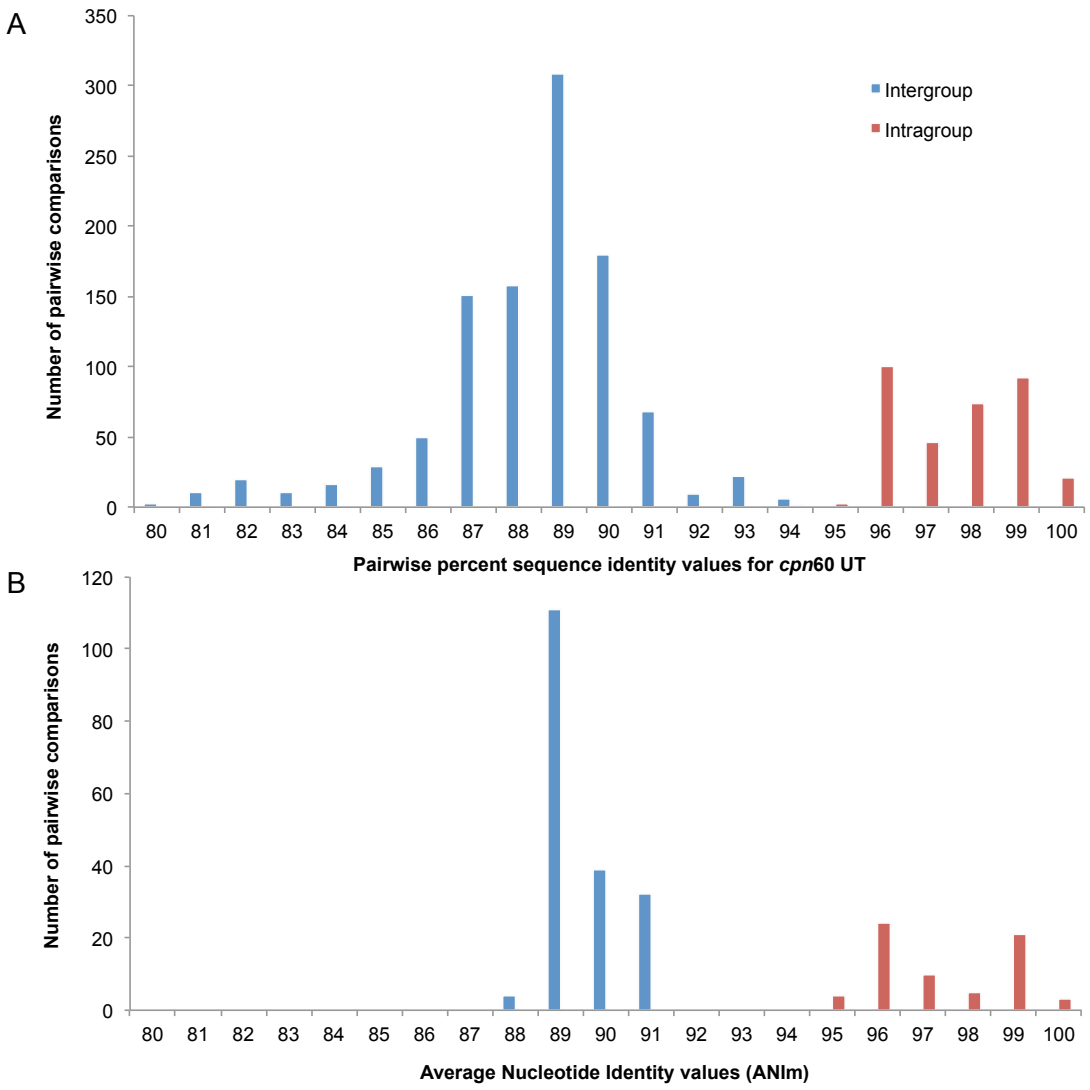


Figure 3.2. Pairwise comparisons of *cpn*60 UT sequences and whole genome ANIm values of *G. vaginalis* isolates.
A: Distribution of *cpn*60 UT sequences pairwise percent identity values for 52 *G. vaginalis* isolates. B: Distribution of pairwise whole genome ANIm values of 23 *G. vaginalis* isolates.

Piot biotyping, sialidase activity and antimicrobial sensitivity

Piot biotyping results for four new study isolates and nine previously described isolates representing the four *cpn*60-defined subgroups are summarized in Table 1. *β*-galactosidase activity was detected only in subgroup C, and these isolates were determined to be Piot biotype 1 or 8. All subgroup B isolates were Piot biotype 5, but subgroup A and C isolates were mixtures of biotypes 2, 5 and 7, and 1 and 8, respectively.

Table 3.2: Biotyping, sialidase and antimicrobial sensitivity properties of study isolates.

*cpn*60 subgroup	Isolate	Piot biotype	Sialidase activity	MIC (µg/mL)[3]	
				Metronidazole	Clindamycin
A	N137	7	-	> 256	0.04
	N134	5	-	> 256	< 0.01
	N158	2	-	> 256	0.04
B	N144	5	+	> 256	0.08
	N170	5	+	> 256	0.05
	N95	5	+	> 256	0.05
	W11	5	+	16	0.10
	N101	5	+	> 256	0.04
	N153	5	+	> 256	0.04
C	N165	8	-	> 256	no growth
	ATCC 14018^T	1	-	3	0.02
	ATCC 49145	8	-	0.38	0.01
D	N160	5	-	> 256	0.02

[1] Piot biotype and sialidase gene detection by PCR of all isolates except N170, N95, W11 and N101 were previously described in Jayaprakash *et al.* (2012) and are included here for context.
[2]L = Lipase, B = *β*-galactosidase, H = Hippurate hydrolase. Biotype designation is according to Piot biotyping scheme (Piot *et al.* 1984).
[3]Minimum inhibitory concentration

MIC values for metronidazole were ≥ 256 μg/mL for all isolates except for W11, ATCC 14018^T and ATCC 49145. Based on National Committee for Clinical Laboratory Standards (NCCLS) guidelines, MIC values for W11, ATCC 14018^T and ATCC 49145 were below the breakpoint of resistance of ≥32 μg/mL for metronidazole. All isolates were sensitive to clindamycin with MIC values ≤ 0.1 μg/mL, below the NCCLS breakpoint of ≥8 μg/ml (Table 3.1). No trend was observed in the relationship between *cpn*60 subgroup affiliation and susceptibility to metronidazole or clindamycin.

All 13 isolates, phenotyped by Piot biotyping (Piot *et al.* 1984) were also tested for sialidase activity using a modified filter spot assay. Interestingly, only subgroup B isolates were positive for sialidase activity in the filter spot assay (Table 3.1). Based on these results, ATCC 14018^T (subgroup C) was selected as negative control and W11 (subgroup B) was selected as positive control in the subsequent experiment. Using a set of previously published primers for detecting the sialidase gene (Santiago *et al.* 2011b), an expanded subgroup of 35 isolates was tested for presence of the sialidase gene. In the expanded group of 35 isolates, all but one of nine new subgroup B isolates was PCR positive, and only one of three new subgroup D isolates was PCR positive. All nine new subgroup C isolates were PCR positive, and only very weak bands were observed for eight of 14 new subgroup A isolates (data not shown). Again, only subgroup B isolates were positive for sialidase activity in the filter spot assay in the expanded set of isolates (Figure 3.3). These findings suggest that enzymatic activity is largely inconsistent with PCR results for identical isolates.

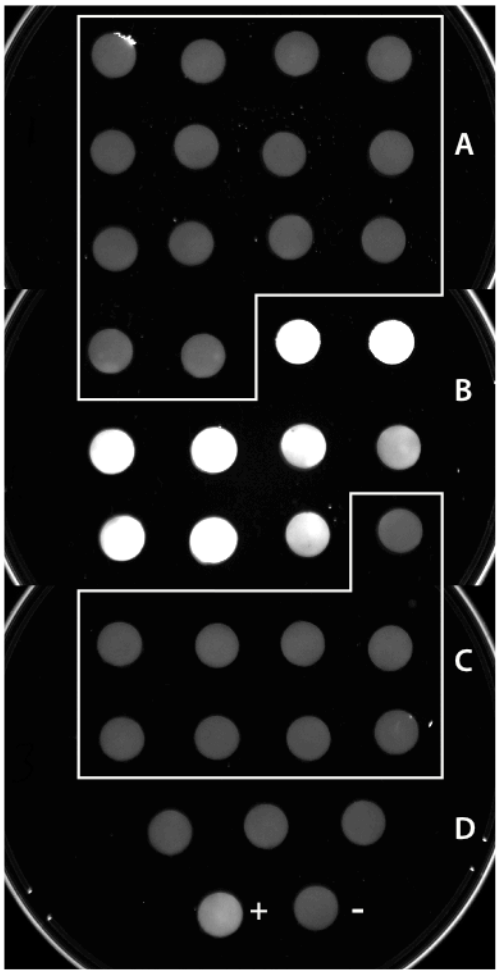


Figure 3.3. Sialidase filter spot test results for *G. vaginalis* isolates.
G. vaginalis isolates (n=35) representing *cpn*60 subgroups A-D were tested using sialidase filter spot assay.W11 and ATCC 14018^{T} are included as positive and negative controls, run on each plate.

Protein profiles

To investigate the relationship of whole cell protein profiles of *G. vaginalis* isolates representing the four subgroups, SDS-PAGE banding patterns of protein extracts from triplicate cultures of 13 isolates were compared. Protein profiles of isolates in each subgroup clustered together, forming distinct clades according to *cpn*60 subgroup (Figure 3.4). Minimum percent similarity values within each subgroup ranged from 58% to 79%. Comparison of SDS-PAGE banding patterns identified an ~40 kDa protein apparently highly expressed by all six subgroup B isolates (Figure 3.5) and an ~45 kDa protein apparently highly expressed by the subgroup D isolate. For a more precise estimate of protein size, extracts of subgroup B isolate N153 and subgroup D isolate N160 were analyzed using the Bioanalyzer Protein 80 chip. The protein bands of interest were found to correspond to apparent protein sizes of 43.7 kDa and 49 kDa respectively (Supplementary Figure 3.1). Protein bands of interest were excised from the SDS-PAGE gels of N153 and N160 for peptide mass spectrometry. For N153, duplicate samples of the protein of interest from separate electrophoresis experiments were collected.

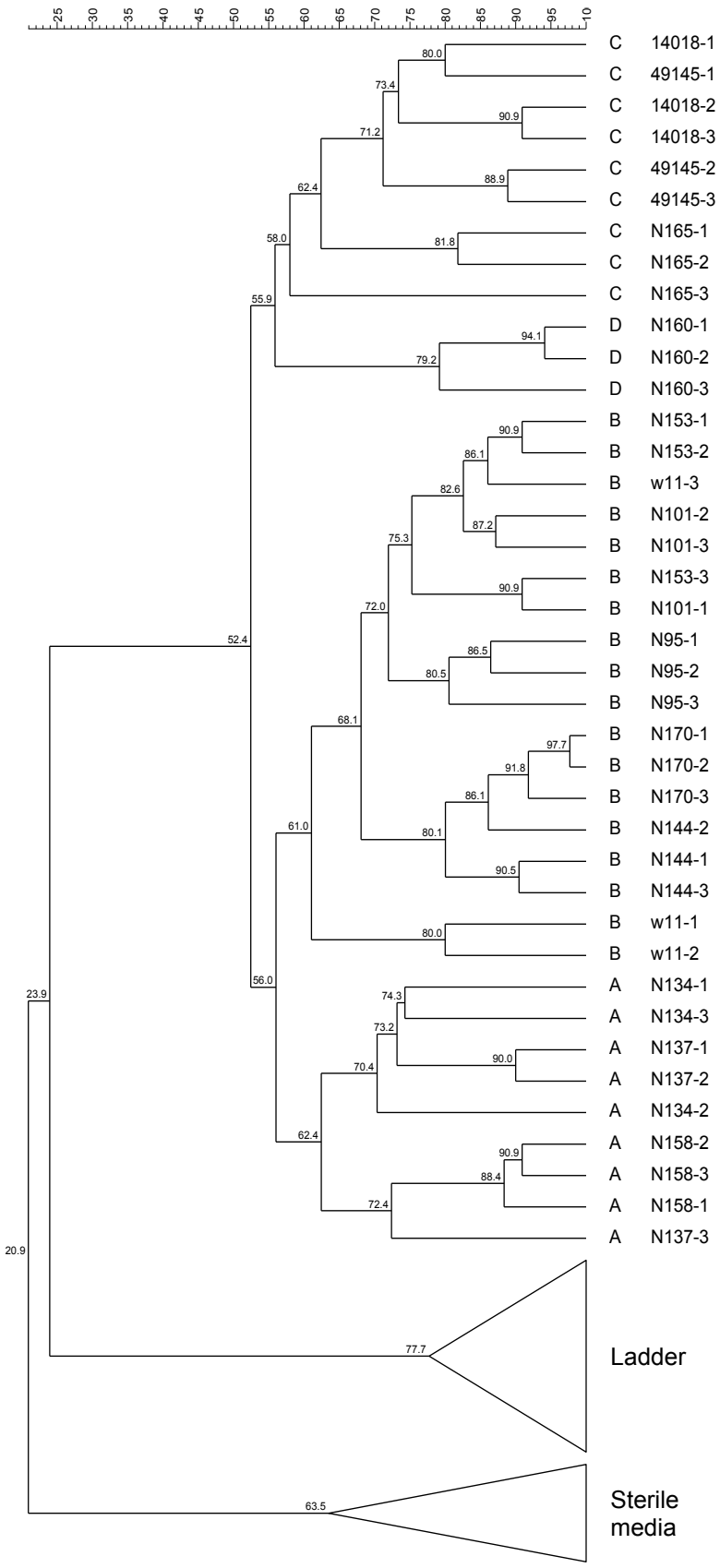

25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 10
C 14018-1
C 49145-1
C 14018-2
C 14018-3
C 49145-2
C 49145-3
C N165-1
C N165-2
C N165-3
D N160-1
D N160-2
D N160-3
B N153-1
B N153-2
B w11-3
B N101-2
B N101-3
B N153-3
B N101-1
B N95-1
B N95-2
B N95-3
B N170-1
B N170-2
B N170-3
B N144-2
B N144-1
B N144-3
B w11-1
B w11-2
A N134-1
A N134-3
A N137-1
A N137-2
A N134-2
A N158-2
A N158-3
A N158-1
A N137-3
Ladder
Sterile media
80.0
73.4
90.9
71.2
88.9
62.4
81.8
58.0
55.9
94.1
79.2
90.9
86.1
82.6
87.2
75.3
90.9
72.0
52.4
86.5
80.5
68.1
97.7
91.8
86.1
61.0
80.1
90.5
80.0
56.0
23.9
74.3
73.2
90.0
70.4
62.4
90.9
88.4
72.4
20.9
77.7
63.5

Figure 3.4. Dendograms of whole cell protein profiles of *G. vaginalis* isolates.
Dendograms were created from whole cell protein profiles using UPGMA clustering of Dice co-efficient values with 0.6% optimization and 0.8% position tolerance values. *cpn*60 subgroup and isolate name, followed by replicate number, are indicated. Protein size marker lanes (Ladder) and sterile media extracts were included as controls for reproducibility of banding patterns across multiple SDS-PAGE gels.

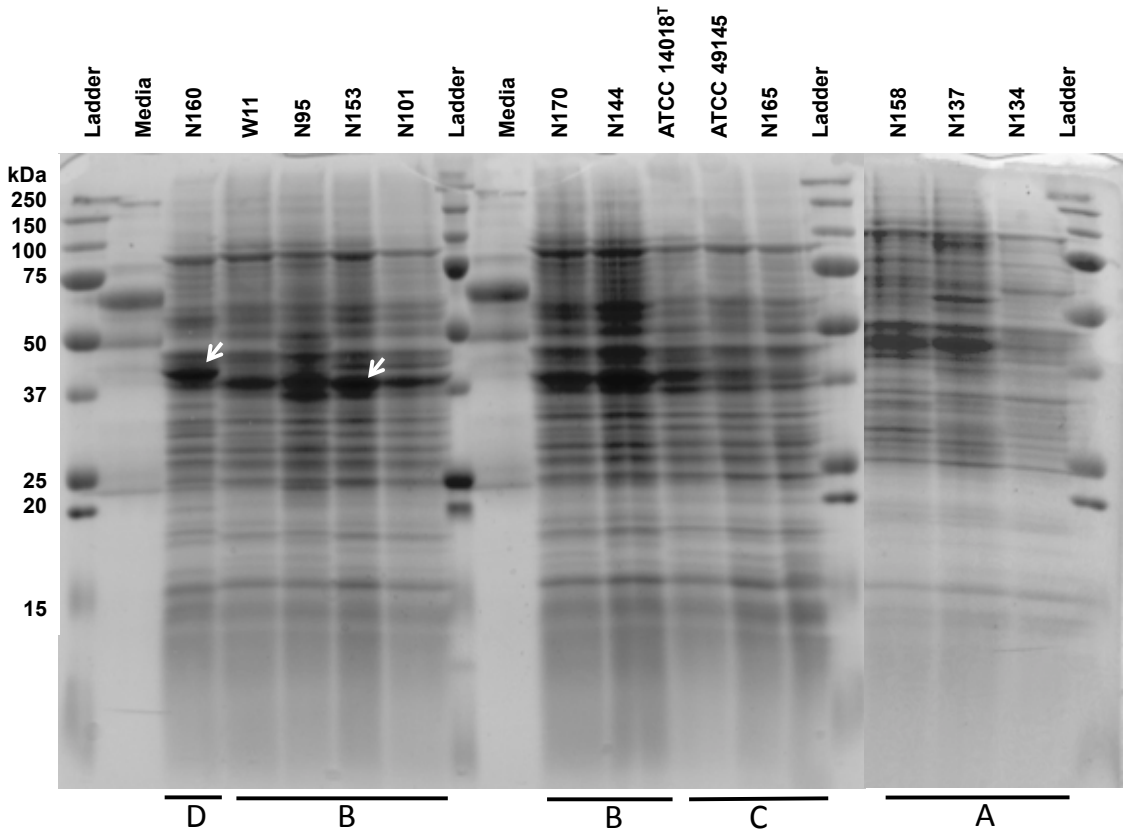
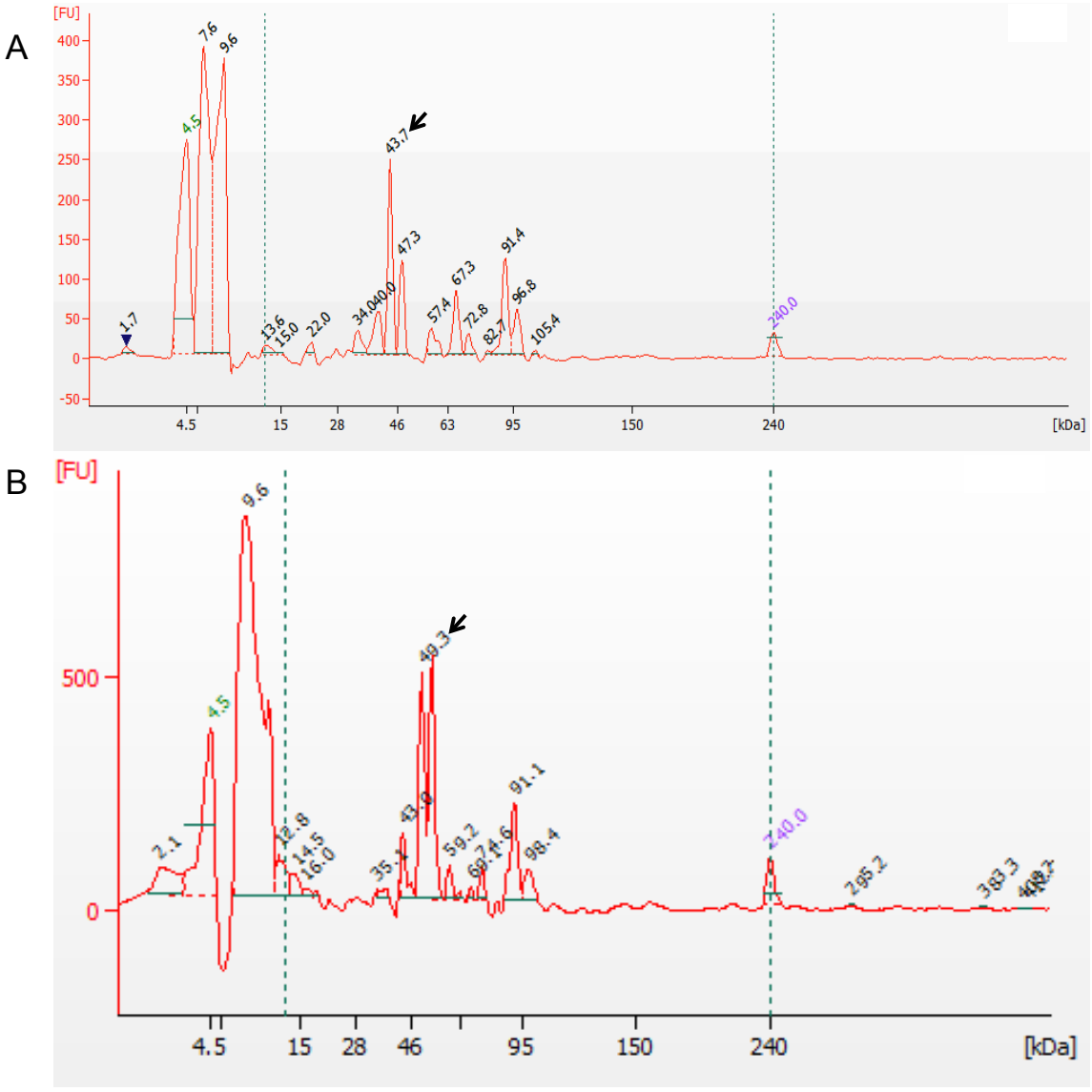


Figure 3.5. Protein profiles of *G. vaginalis* isolates.
Protein extracts (75 µg) from isolates of *G. vaginalis* from four *cpn*60 subgroups (indicated below gel) and sterile broth (Media) were run on SDS PAGE (12%) along with 15 µg of ladder. Bands excised for peptide mass fingerprinting are indicated with white arrows.



Supplementary Figure 3.1. Bioanalyzer profiles of protein extracts from *G. vaginalis* isolates
A: Bioanalyzer profile for protein extract from *G. vaginalis* N153 showing peak at 43.7 kDa (arrow). B: Bioanalyzer profile for protein extract from *G. vaginalis* N160 showing peak at 49.3 kDa (arrow).

In both replicate experiments with N153, peptide mass patterns matched a predicted ORF in the N153 genome (Gene ID in JGI: 2528387946, encoding 409 aa,) that was annotated as an ABC transporter substrate-binding protein belonging to COG2182 (maltose-binding periplasmic proteins/domains). Since the whole genome sequence of subgroup D isolate N160 was not available, peptides from N160 were compared to the predicted proteome of *G. vaginalis* B479 (1500E) (NZ_ADES00000000.1), which was identified as belonging to subgroup D based on *cpn*60 UT sequence analysis (Figure 3.1). The protein excised from extracts of N160 corresponded to an ORF in *G. vaginalis* B479 (1500E) (EIK82014.1, encoding 448 aa), which was annotated as an ABC sugar transport system periplasmic component belonging to COG1653 (ABC-type sugar transport system, periplasmic component).

The primary sequence of proteins identified in subgroup B and subgroup D isolates were searched for conserved domains using a NCBI conserved domain search against the conserved domain database (CDD). Two domains were identified within the protein identified from the subgroup B isolate N153 genome: periplasmic binding component of ABC transport system specific for an unknown oligosaccharide (periplasmic binding fold, PBP2_oligosaccharide_1) and a malE type maltose-binding periplasmic domain. Four conserved domains were identified the predicted protein from subgroup D isolate *G. vaginalis* B479 (1500E): a bacterial extracellular solute-binding protein domain (SBP_bac_8); type 2 periplasmic binding fold superfamily (Periplasmic_Binding_Protein_Type_2 super family); ABC-type sugar transport system, periplasmic component (UgpB), and magnesium chelatase domain, H subunit (BchH).

To investigate the presence of orthologous sequences in other *G. vaginalis* isolates corresponding to the proteins of interest, genomes of *G. vaginalis* from representatives of subgroups A, B, C and D (35 total genomes) were searched using BLASTp with the proteins of interest as queries. Sequences of predicted protein orthologs and their corresponding nucleotide sequences were extracted from the annotated genomes. Orthologous nucleotide sequences from *Bifidobacterium* species were identified using the IMG/ER database (Joint Genome Institute, https://img.jgi.doe.gov/cgi-bin/er/main.cgi) and included in the subsequent phylogenetic analysis. Phylogenetic trees showing the relationships of genes predicted to encode subgroup- specific proteins and their nearest neighbours in other subgroups are shown in Figure 3.6 and Figure 3.7.

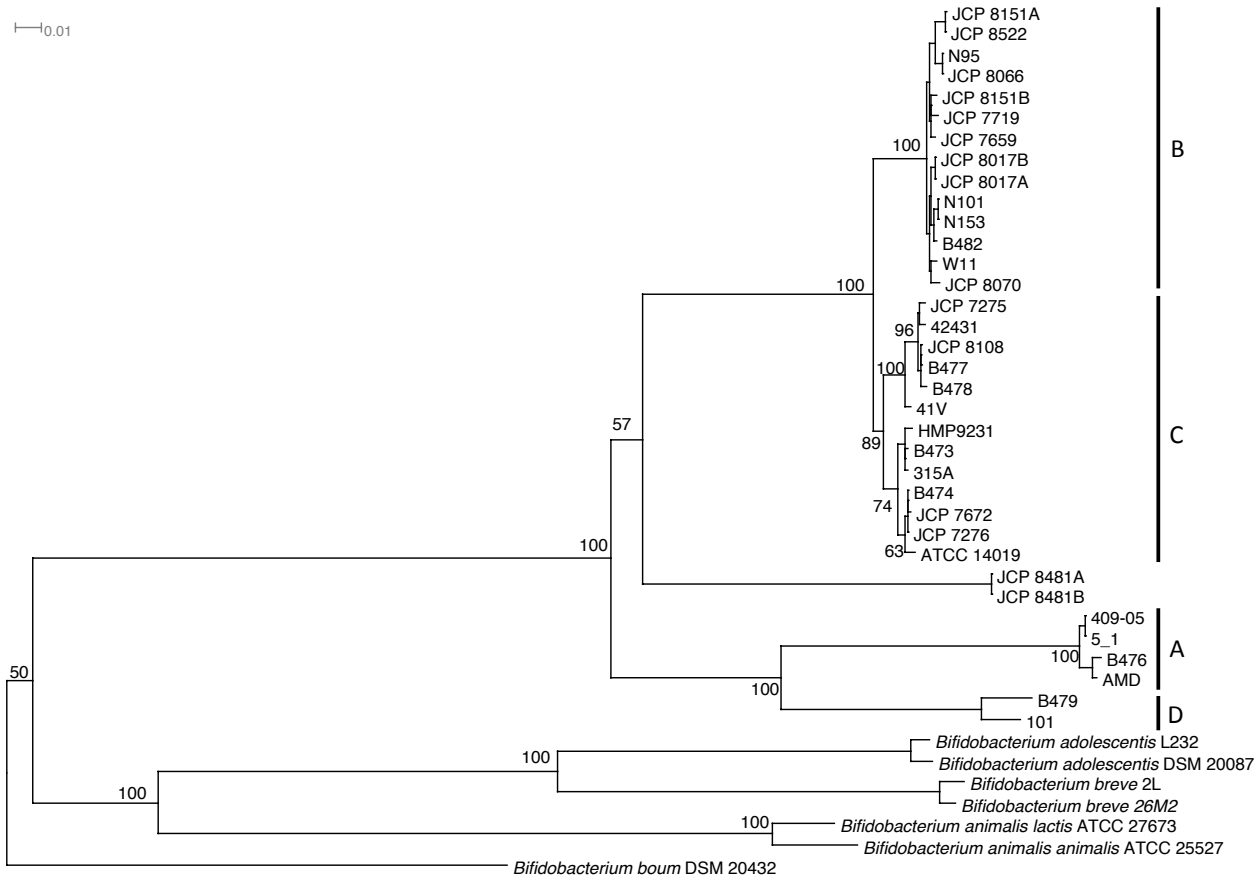


Figure 3.6. Phylogenetic tree of nucleotide sequences predicted to be ABC transporter substrate-binding protein (subgroup B) from *G. vaginalis* and *Bifidobacterium* spp.
Phylogenetic tree of nucleotide sequences (1295 bp alignment) of best matches from *G. vaginalis* genomes (n=35) from four subgroups (A – D) and from genomes of *Bifidobacterium* spp. to an ORF encoding a predicted ABC transporter substrate-binding protein identified in subgroup B isolate N153.

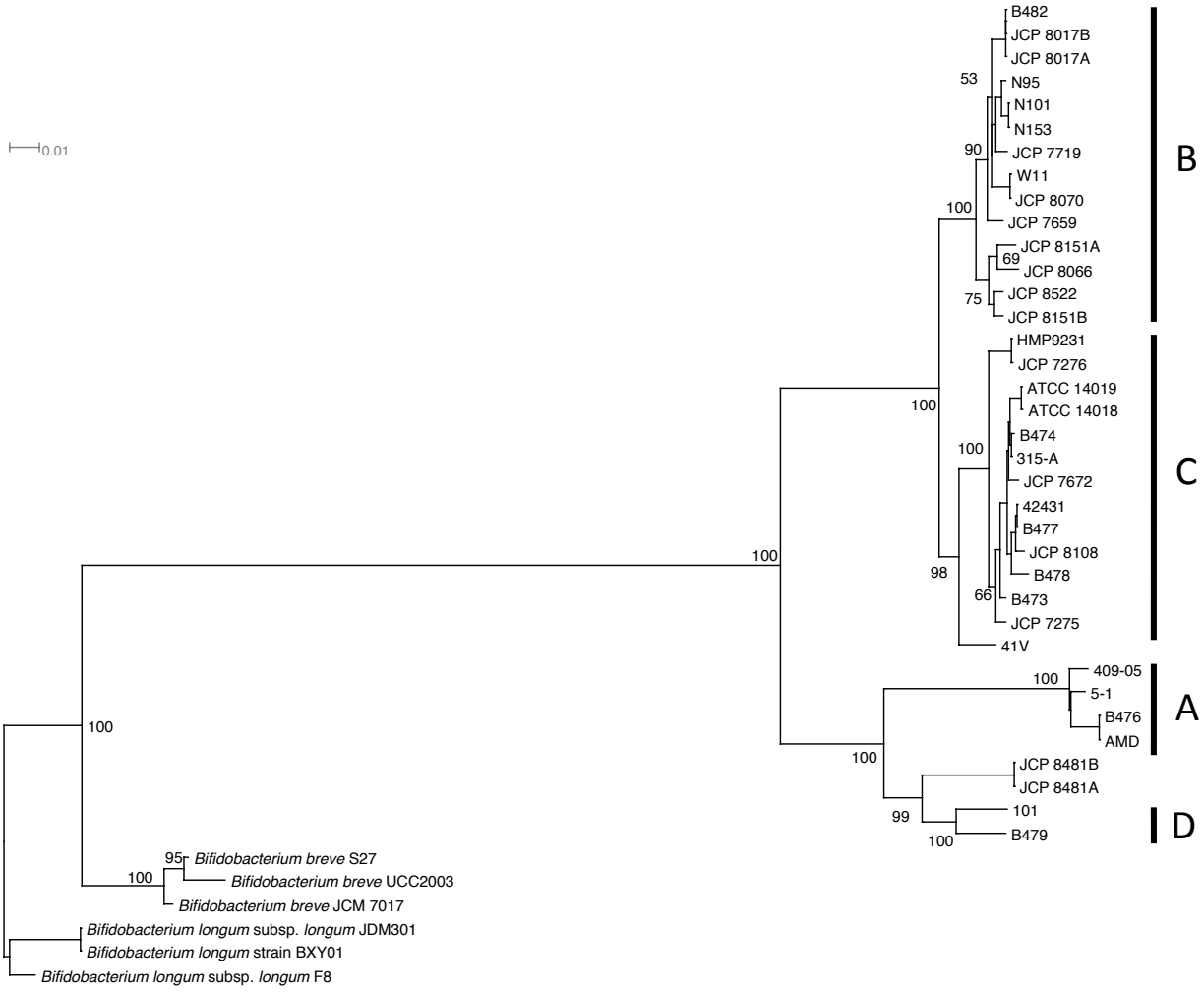


Figure 3.7. Phylogenetic tree of nucleotide sequences predicted to be ABC transporter substrate-binding protein (subgroup D) from *G. vaginalis* and *Bifidobacterium* spp.
Phylogenetic tree of nucleotide sequences (1389 bp alignment) of best matches from *G. vaginalis* genomes (n=36) from four subgroups (A – D) and from *Bifidobacterium* species to an ORF encoding a predicted ABC transporter substrate-binding protein identified in subgroup D isolate N160.

An ortholog of the predicted maltose-binding protein encoding gene identified in N153 was identified in all subgroup B genomes, with nucleotide sequence identities >98% to N153 (Figure 3.6). Sequences with 94-95% nucleotide sequence identity were identified in subgroup C genomes, but these sequences formed a distinct clade in the tree with good bootstrap support. The subgroup B and C orthologs were also distinguishable at the protein sequence level (data not shown). Nucleotide sequence identities between subgroup B and related sequences from subgroups A and D were much lower (<78% nucleotide sequence identity), and similar to those observed between *Bifidobacterium* species. These findings indicate that different protein profiles of the *G. vaginalis* subgroups may result from the expression of a gene that is either present (in subgroup B) or absent (subgroups A, C and D), rather than differential expression of a conserved gene.

A similar clustering pattern was observed for the putative ABC transporter substrate-binding protein identified in subgroup D (Figure 3.7). A closely related sequence (94% identity) with a common node with good bootstrap support was identified in the only other subgroup D isolate (101) for which a whole genome sequence is available. These subgroup D specific sequences were clearly distinct from the nearest sequences in the other subgroups (85-89% identical to orthologous sequences in subgroups A, B and C).

3.5 Discussion

Previous investigations of the relationship of *G. vaginalis* biotypes or genotypes and women's clinical characteristics have been limited by dependence on diagnostic features that require culture of isolates. Methods such as biotyping (Benito *et al.* 1986; Piot *et al.* 1984) and ARDRA (Santiago *et al.* 2011b) require culture isolation before they

can be applied. The association of any single *G. vaginalis* isolate with a complex clinical entity such as BV is problematic since *G. vaginalis* can readily be isolated from clinically healthy women as well as those with BV (Aroutcheva *et al.* 2001; Verhelst *et al.* 2004), and individual women can be colonized by a diverse mixture of *G. vaginalis* strains simultaneously (Aroutcheva *et al.* 2001; Chaban *et al.* 2014; Paramel Jayaprakash *et al.* 2012). In other words, the observation that a particular isolate is from a woman diagnosed with BV is not sufficient evidence to infer causation or designate that isolate as "pathogenic". Understanding the role of diverse *G. vaginalis* strains in vaginal microbial ecology and determining any systematic association with BV and other health outcomes depends upon the further elucidation of characteristics that differentiate *G. vaginalis* strains, as well as the development of methods to detect biologically relevant groups in the context of the entire microbiome.

Our phylogenetic analysis confirms that four recently described “clades” (Ahmed *et al.* 2012) are the same as the four *cpn*60-defined subgroups defined in our previous work (Paramel Jayaprakash *et al.* 2012) (Figure 1). Furthermore, the bimodal distribution of pairwise ANIm values is consistent with the suggested bacterial species level cut-off of 95-96% proposed by Richter *et al.* (2009), with all inter-subgroup identities ≤91% and all intra-subgroup identities ≥95% (Figure 3.2). The consistency between these reported genotyping schemes was not unexpected given previous demonstrations of the prediction of whole genome similarity based on *cpn*60 universal target sequences (Verbeke *et al.* 2011). Although the genotypic requirements for designation of the four *G. vaginalis* subgroups as separate species have arguably been met (Richter and Rossello-Mora 2009),

identification of additional phenotypic properties that consistently resolve these subgroups is required before a formal proposal of reclassification can be made.

Antimicrobial resistance patterns were not helpful in subgroup resolution. According to the National Committee for Clinical Laboratory Standards (NCCLS) guidelines, the breakpoints of resistance are ≥32 μg/mL for metronidazole and ≥8 μg/mL for clindamycin. In the current study, all isolates other than the ATCC strains were resistant to metronidazole, but all isolates tested were also sensitive to clindamycin. These results are consistent with previous reports (Goldstein *et al.* 2002; Shanker *et al.* 1982; Tomusiak *et al.* 2011). Metronidazole and clindamycin are the two main drugs used by clinicians for the treatment of BV (Greaves *et al.* 1988; Powell 2013) and are administered either orally or compounded in gels or creams. There are however numerous reports of resistance to metronidazole by *G. vaginalis* isolates (Aroutcheva *et al.* 2001; Goldstein *et al.* 2002; Kharsany *et al.* 1993; McLean and McGroarty 1996). Proposed mechanisms of metronidazole resistance in *G. vaginalis* include inactivation or deletion of genes encoding nitroreductase activity (Edwards 1993), biofilm formation (Swidsinski *et al.* 2008), and inhibition of penetration of the drug by a dense layer of fimbriae on the cell surface (McLean and McGroarty 1996).

Although there was no clear pattern of Piot biotypes among the representative isolates, the detection of *β*-galactosidase activity was unique to subgroup C, and may represent a potential defining characteristic for this subgroup (Table 3.1). All isolates from subgroup B belong to Piot biotype 5 and are positive for sialidase enzyme activity (Table 3.1 and Figure 3.3), while isolates from other subgroups belonged to a number of different biotypes and were all-negative for sialidase activity. In a recent study where

sialidase activity was determined quantitatively (Lewis *et al.* 2013), a wide range of activity levels was reported among isolates. Our analysis of *cpn*60 sequences from isolates included in this previous study (Figure 1) shows that all but one sialidase positive strain belong to *cpn*60 subgroup B (isolates with names beginning with JCP in Figure 3.1). The one remaining sialidase positive isolate (JCP7276) belongs to subgroup C. Although all isolates of subgroup C included in the current study were PCR positive for sialidase gene, none had detectable sialidase enzyme activity. Taken together, these findings may indicate that some subgroup C isolates have sialidase activity at levels below the detection limit of the filter spot assay used in this study. Differential sialidase production by *G. vaginalis* subgroups has potential clinical significance since sialidase activity of vaginal fluid is proposed as a marker for diagnosing BV (Briselden *et al.* 1992; Cauci *et al.* 2005; Shujatullah *et al.* 2010). Sialidase and other mucolytic enzymatic activities in vaginal fluid are likely detrimental to the protective mucous layer and have been proposed to play a role in recurrent BV (McGregor *et al.* 1994).

Total protein profiling using PAGE is an important molecular technique used for bacterial systematics, and for species or strain level resolution (Hesselberg and Vreeland 1995; Jackman 1982; Stanley *et al.* 1993). The four subgroups of *G. vaginalis* were resolved by protein profiles in this study. For subgroups B and D, we identified putative biomarker proteins, both of which are predicted to be components of ATP-binding cassette (ABC) transporters. Similar proteins are present in all bacteria, where they participate in transporting a variety of molecules, including nutrients, across the cellular membrane (Higgins 1992, 2001). ABC transporters include a substrate binding protein, two transmembrane proteins, and two membrane-associated nucleotide-binding proteins.

Based on their primary amino acid sequences, both subgroup B and D biomarker proteins were identified as having domain characteristics consistent with extracellular solute binding proteins, and more specifically, domains characteristic of maltose/maltodextrin-binding proteins typified by MalE in *E. coli* (Ehrmann *et al.* 1998), and *sn*-glycerol-3-phosphate (G3P) binding protein UgpB in *E. coli* (Overduin *et al.* 1988), respectively. Similarity between the Mal and Ugp ABC transporter systems has been noted (Hekstra and Tommassen 1993), and the identification of proteins that may be involved in uptake of maltose/maltodextrin in *G. vaginalis* is noteworthy given the significance of these nutrients in the vaginal ecosystem. Glycogen, expressed by vaginal epithelium, promotes the colonization of *Lactobacillus*. However, most vaginal isolates of *Lactobacillus* cannot degrade glycogen and utilize it directly (Stewart-Tull 1964; Wylie and Henderson 1969). Instead, *Lactobacillus* utilizes the breakdown products of glycogen including maltose, maltotriose, maltopentaose, and maltodextrin, which are products of human α-amylase activity on glycogen (Spear *et al.* 2014). Efficient uptake of glycogen breakdown products by some *G. vaginalis* strains, facilitated by high levels of expression of ABC transporter components, would lead to direct competition with *Lactobacillus*, which could have significant consequences for vaginal microbial community composition and dynamics. Although fermentation of maltose is a diagnostic characteristic of *G. vaginalis* (Greenwood and Pickett 1979), the mechanisms for its uptake or the uptake of related maltodextrin compounds are not known. Further *in vitro* characterization of the putative biomarker proteins identified in this study to elucidate substrate specificity and expression patterns is ongoing.

In summary, our results demonstrate that "clades" based on supragenome genome analyses of 17 *G. vaginalis* genomes, correspond to *cpn*60 subgroups, and furthermore that *cpn*60 subgroups are distinguishable based on whole cell protein profiles. The detection of sialidase enzymatic activity in only one of the *cpn*60-defined subgroups and the identification of subgroup-specific proteins that likely play a role in nutrient uptake support the hypothesis that *G. vaginalis* subgroups are also ecologically distinct, having different relationships with the host and the other constituents of the vaginal microbial community.

CHAPTER 4 - Identification and characterization of anaerobic and facultative isolates of *Gardnerella vaginalis* subgroup B.

Citation

Jayaprakash TP, Schellenberg JJ, Hill JE. Identification and characterization of anaerobic and facultative isolates of *Gardnerella vaginalis* subgroup B.

Author Contributions

Conceived and designed the experiments: TPJ JJS JEH. Performed the experiments: TPJ. Analyzed the data: TPJ JJS JEH. Wrote the paper: TPJ JJS JEH.

Fourth Chapter Transition

This chapter follows observations made during investigation of objectives accomplished in previous chapters, one of which was the observation of the potential clinical significance for subgroup B *G. vaginalis*. Subgroup B *G. vaginalis* was found to be significantly more abundant in a cohort of Kenyan women grouped into 'BV' category than 'Normal' based on the Nugent score. Subgroup B was the only group with sialidase enzymatic activity among the four subgroups of *G. vaginalis*. It was also observed that two isolates belonging to subgroup B were obligate anaerobes that did not grow in 7% CO_2 atmosphere.

Due to the significant relationship illustrated throughout the literature between anaerobes and BV, anaerobic isolates of *G. vaginalis* were of substantial interest. Hence, this chapter focuses on addressing how genomes of anaerobic isolates of subgroup B differ from facultative isolates of subgroup B *G. vaginalis.*

4.1 Abstract

Four distinct subgroups of the vaginal bacterium *Gardnerella vaginalis* have been defined based on *cpn*60 universal target sequences. Sialidase enzymatic activity is confined to subgroup B and this subgroup has previously been observed to be significantly more abundant in the vaginal microbiomes of African women with BV than those with healthy vaginal microbiota. In the current study, the growth characteristics of subgroup B isolates in an anaerobic atmosphere and in a 7% CO_2 atmosphere were investigated. While all isolates grew anaerobically, some isolates were incapable of growth in 7% CO_2. To determine genome content differences that could account for this phenotypic difference, the whole genome sequence of four *G. vaginalis* subgroup B isolates representing facultative and anaerobic phenotypes were determined. A 15 kb contiguous fragment present in the facultative isolates W11 and N95, as well as in 703C2 and ATCC 14109 (publically available genome sequences, reported to grow in 7% CO_2), was absent in anaerobic isolates N153 and N101. The fragment contains genes predicted to encode an epimerase, reductase and ABC transporter that might be involved in protecting the cells from oxidative damage and allowing the cells to grow in the presence of oxygen. Similarly, a 21 kb contiguous fragment was present only in the obligate anaerobic strains, which based on predicted functionality of COG categories for ORFs in the fragment, are speculated to be involved in cell wall biogenesis. Given the well-known relationship between an anaerobic microbiota and BV, anaerobic isolates of *G. vaginalis* are potentially important players in the vaginal microbial community.

4.2 Introduction

The clinical significance of *G. vaginalis* was described even before its current name was proposed. It was initially associated with the condition of non-specific vaginitis in cases where other agents of vaginitis such as *Trichomonas* and *Candida* were not detected (Catlin 1992). Diversity within *G. vaginalis* is well recognized, and recently genomic and phenotypic evidence has been provided to demonstrate that *G. vaginalis* comprises at least four different subgroups that likely represent different species. Differential distribution of the subgroups among women with varying clinical status (Balashov *et al.* 2014; Paramel Jayaprakash *et al.* 2012) suggests that *G. vaginalis* should not be considered as a single entity in the evaluation of vaginal health.

Sialidase activity is considered an important virulence factor in BV and although a gene predicted to encode sialidase can be detected by PCR in both subgroup B and C, only subgroup B isolates have been shown to have sialidase enzyme activity using the filter spot test ((Paramel Jayaprakash *et al.* 2012) and Chapter 3). Another indication of the potential clinical significance of *G. vaginalis* subgroup B comes from the observation that this subgroup was significantly more abundant in samples from women with BV than those without BV in a study of 44 Kenyan women (Paramel Jayaprakash *et al.* 2012). Our interest in further characterization of the *G. vaginalis* subgroups led to growth experiments in which it was observed that all isolates from all subgroups grew on Columbia blood agar containing 5% sheep blood in a 7% CO_2 atmosphere except for one isolate from subgroup B. The isolate from subgroup B N153 did not grow even after 84 hrs. However, when isolates were incubated with anaerobic gaspaks, growth was observed for all isolates within 48 hrs.

G. vaginalis has been proposed to be an initial colonizer during the transition from a healthy vaginal microbiome to BV (Patterson *et al.* 2010; Schellenberg *et al.* 2011b). Anaerobic *G. vaginalis* strains may play a role in establishing a BV environment, producing biofilm and acting as stage setters for the takeover of strict anaerobes, ultimately leading to the shift from lactobacilli to a mixed anaerobic microbial community that defines BV. The isolation of obligate anaerobic strains of *G. vaginalis* has been reported occasionally (Catlin 1992; Malone *et al.* 1975), but no further characterization of these isolates has been done and the genomic determinants responsible for this phenotypic property are unknown.

The goals of the current study were to investigate growth of isolates of *G. vaginalis* subgroup B in anaerobic and aerobic environments. Based on the initial results of this investigation that illustrated differences in tolerance of oxygen, we compared the whole genome sequences of anaerobic and facultative subgroup B isolates to identify gene content differences that might explain the observed phenotypic differences.

4.3 Materials and Methods

Isolates and culture conditions

G. vaginalis isolates were cultured on Columbia blood agar plates containing 5% sheep blood (BD, Mississauga, ON) or BBL™ V Agar (BD, Mississauga, ON), which is selective media for *G. vaginalis* containing 5% human blood. For resurrection of archived isolates, Columbia sheep blood agar plates were inoculated from -80 °C frozen stocks and incubated at 37 °C for 72 hrs under anaerobic conditions generated using the GasPak EZ Anaerobe Pouch System (BD, Mississauga, ON), which produces an atmosphere with

only residual oxygen (<0.1%) and ≥10% CO_2 within 90 minutes. Isolates were subcultured once and their identities confirmed prior to the growth experiment by amplifying and sequencing the *cpn*60 universal target as previously described (Hill *et al.* 2006a). Amplicon products were purified (EZ-10 spin column PCR product purification kit, Bio Basic Canada Inc., Markham, ON, Canada) and sequenced using Sanger sequencing. Raw sequence data were assembled using pregap4 and gap4 (Staden *et al.* 2000) and compared to the cpnDB database (www.cpndb.ca) for identification.

For growth comparisons in different atmospheres, streak-based culture plates were made using a loopful (1µl loop, Fisher Scientific, Catalog No.: 22-363-603, Whitby, ON, Canada) of culture for each isolate. Plates were incubated at 37 °C for 48 hrs under two sets of conditions: in a 7% CO_2 incubator or anaerobic conditions as described above. Growth was scored on a scale of 0 to 4 where a score of 0 indicates no colonies, a score of 1 indicates colonies on the first streak and <10 colonies on the second streak, a score of 2 indicates ≥10 colonies on the second streak and <10 colonies on the third streak, a score of 3 indicates ≥10 colonies on the third streak and <10 colonies on the fourth streak and a score of 4 indicates ≥10 colonies on the fourth streak. Triplicate experiments were performed with two plates evaluated per isolate in each triplicate (six plates total per isolate). Results for each isolate were reported as an average of the six individual scores.

Genomic DNA extraction and whole genome sequencing

G. vaginalis isolates N153, N101, W11 and N95 were grown in ATCC broth #1685 anaerobically at 37 °C to stationary phase, and genomic DNA was extracted from cultures using a modified salting out procedure (Martín-Platero *et al.* 2007). Extracts

were stored at -80 °C in TE buffer (1 mM EDTA, 10 mM Tris HCl, pH 7.4). Identification of isolates was confirmed by amplification and sequencing of the *cpn*60 universal target (UT) region (Hill *et al.* 2006a). The quality of genomic DNA extractions was assessed by spectrophotometric measurements (A_{260}:A_{280} ratio of 1.8-2.0).

Sample preparation and whole genome sequencing was done as described in the GS Junior whole genome shotgun sequencing protocol (Roche, Brandford, CT). Sequence assembly was done using gsAssembler (v2.7).

Genome sequences were submitted to the Joint Genome Institute (JGI - https://img.jgi.doe.gov/cgi-bin/er/main.cgi) for annotation service. Publically available whole genome sequence for *G. vaginalis* B482 (00703C2mash, Genbank Accession NZ_ADEU00000000) was also used for the whole genome comparison as this strain has a *cpn*60 sequence identical to all four sequenced *G. vaginalis* isolates, and is described in the literature as being able to grow in 6% CO_2 (Ahmed *et al.* 2012).

COG categories were assigned to the open reading frames of individual genomes by JGI. Average Nucleotide Identity (ANI) between whole genome sequences was calculated using the MUMmer algorithm (ANIm) within JSpecies (Richter and Rossello-Mora 2009). Pangenome analysis of *G. vaginalis* genomes was conducted using *GView* (Petkau *et al.* 2010) server for CDS (coding DNA sequence) with the following settings: genomes were compared against each other using tblastn with expect value cutoff of 1×10^{-10}; coding sequences (CDS) from genomes were translated into amino acid sequence using the genetic code of Bacteria; minimum alignment length cutoff value of 100 and minimum percentage identity cutoff value of 70%.

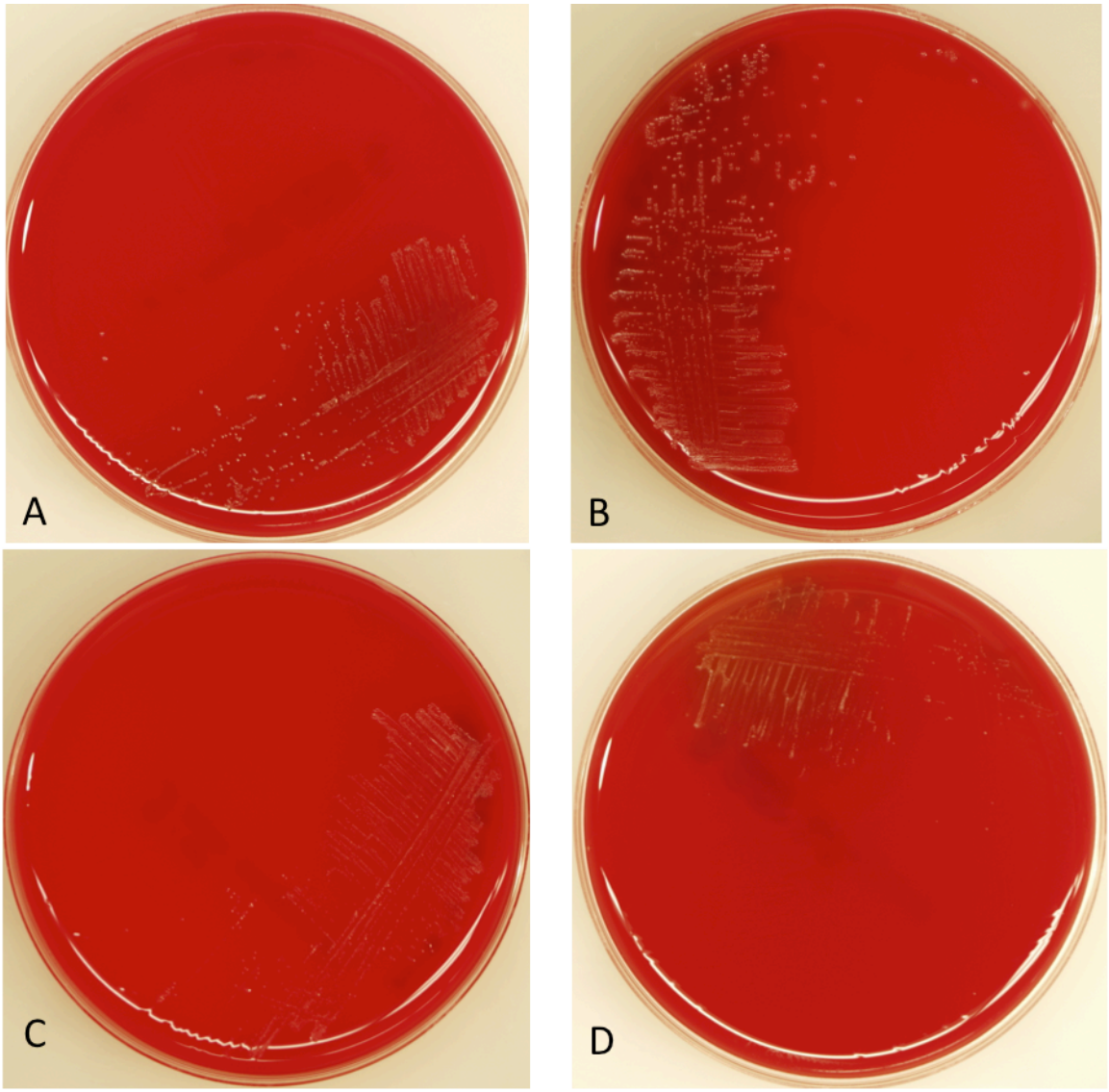
4.4 Results

Following the observation of a single isolate (N153) of subgroup B not being able to grow in 7% CO_2 atmospheric conditions during a growth curve experiment, all available isolates of subgroup B (N153, N144, N170, N101, N95 and W11) along with the type strain ATCC 14018 (subgroup C) were tested for growth under four sets of conditions: 7% CO_2 or anaerobic atmosphere on either Columbia sheep blood agar or human blood based media (V agar). All isolates formed colonies on Columbia sheep blood agar plates as well as V agar when incubated anaerobically at 37 °C for 48 hrs, with scores ranging from 2 to 3 (Table 4.1). All isolates except N153 and N101 formed colonies in 7% CO_2 after 48 hrs. N153 and N101 did neither grow on Columbia sheep blood agar plates nor on V agar even after 72 hrs when incubated in 7% CO_2 atmospheric conditions. For N170, growth was observed with average scores of 1.8 to 2.5 under all conditions, but the colonies formed on either medium in the 7% CO_2 atmosphere were much smaller in comparison to colonies formed on plates incubated anaerobically (Table 4.1 and Supplementary Figure 4.1).

Table 4.1. Growth of *G. vaginalis* subgroup B isolates and type strain on sheep blood agar and human blood agar (V agar) in anaerobic or 7% CO_2 atmospheres. Scores shown are an average ± standard deviation of 6 replicate plates (3 biological replicates and 2 technical replicate for each biological replicate run).

		Sheep blood agar		V agar	
Isolate	**Subgroup**	**Anaerobic**	**7% CO_2**	**Anaerobic**	**7% CO_2**
N144	B	2.50 ± 0.55	2.17 ± 0.98	2.50 ± 0.84	2.33 ± 0.52
N153		2.17 ± 0.75	0.00 ± 0.00	2.33 ± 0.52	0.00 ± 0.00
N170		2.50 ± 0.55	1.83* ± 0.75	2.17 ± 0.63	2.00* ± 0.89
N101		2.17 ± 0.41	0.00 ± 0.00	2.50 ± 0.55	0.00 ± 0.00
N95		2.67 ± 0.52	3.00 ± 0.00	3.00 ± 0.00	2.67 ± 0.52
W11		2.50 ± 0.82	2.17 ± 0.41	2.00 ± 0.00	2.00 ± 0.00
ATCC 14018^T	C	2.67 ± 0.52	2.83 ± 0.41	2.83 ± 0.41	2.83 ± 0.41

*Tiny colonies



Supplementary Figure 4.1: *G. vaginalis* N170 on human blood based agar plate (V agar) and on Columbia blood agar plates containing sheep blood

A: N170 on human blood based agar plate (V agar) incubated at anaerobic conditions. B: N170 on Columbia blood agar plates containing sheep blood incubated at anaerobic conditions. C: N170 on human blood based agar plate (V agar) incubated at 7% CO_2. D: N170 on Columbia blood agar plates containing sheep blood incubated at 7% CO_2.

To determine genome content differences between obligate anaerobic and facultative isolates, N153, N101, W11 and N95 were chosen for whole genome analysis. These isolates were chosen since their *cpn*60 UT sequences are identical, and *cpn*60 UT sequence relationships have been shown to predict whole genome sequence relationships (Verbeke *et al.* 2011). Based on this, we predicted that genome sequences of anaerobic isolates N153 and N101 and facultative isolates W11 and N95 would all be very similar to each other, and that any gene content differences identified would be likely to be responsible for the anaerobic versus facultative phenotypes. Though N153 and N101 were isolated from the same woman, they were chosen for whole genome sequence comparisons for two reasons. First, these two isolates were the only ones available that exhibited the strict anaerobic phenotype. The second reason was that it is known that women can be colonized by multiple strains of *G. vaginalis* (Schellenberg *et al.* 2012), and these two isolates might be different strains with gene content difference that can be ruled out as being responsible for the anaerobic phenotype.

Whole genome shotgun sequencing of the four selected isolates was performed, resulting in 152,896 to 194,763 high quality filter passed reads per genome with an average read length of 463, and providing 39× to 59× coverage (Supplementary Table 4.1). Reads from each isolate were assembled using the Newbler assembly algorithm of the Roche de novo assembler (v2.7), resulting in 7-10 large contigs (>500 bp) per genome and N50 values (defined as the size of the smallest contig in the list of contigs (with equal or larger size) that will amount to half of the assembled genome) of 485,233 (N153), 689,064 (N101), 450,244 (W11) and 386,601 (N95). Since *G. vaginalis* B482 has an identical *cpn*60 sequence to the four study isolates and is reported to grow in 6%

CO_2 (Ahmed *et al.* 2012), the draft genome sequence of B482 (NZ_ADEU00000000) was included in the subsequent analysis. B482 is a clinical isolate from a woman with a Nugent score of 10 and was obtained from the Magee-Women's Research Institute, Pittsburgh, PA (Ahmed *et al.* 2012).

Supplementary Table 4.1: Comparison of genome sequencing run parameters of *G. vaginalis* genomes N153, N101, N95 and W11.

	Benchmark	**N153**	**N101**	**N95**	**W11**
Raw wells	≤250,000	197,876	239, 283	245,687	255,250
Keypass wells	>90%	191,816	233,283	239,100	248,817
High Quality reads	50,000-80,000	152,896	194,763	162,783	162,913
% Filter passed	>50%	79.71	83.39	68.08	63.8
% Mixed+Dots	<20%	3.3	2.73	4.13	4.2
% Too Short Quality	<30%	16.7	16.54	23.36	28.93
% Too Short Primer	<2%	0.009	0.013	0.0002	0.00009

GenBank format genome sequences of N153, N101, W11, N95 and B482 were submitted to the Joint Genome Institute (https://img.jgi.doe.gov/cgi-bin/er/main.cgi) and were annotated using the JGI Integrated Microbial Genomes Expert Review (IMG/ER) service. A description of genome features for *G. vaginalis* isolates N153, N101, W11, N95 and B482 is given in Table 4.2. All four sequenced *Gardnerella* genomes (N153, N101, W11 and N95) and B482 were ~1.5 Mb in size with 42% G+C content. All five genomes have 45 tRNA genes and one copy of the 16S rRNA gene, indicating that if there are multiple copies of the16S rRNA operon, they are either very similar or identical to the assembled one. Gene, COG, Pfam and KEGG count for these four genomes were 1234-1261, 959-973, 1009-1038 and 401-409 respectively. There was no obvious pattern in variation of the parameters described in Table 4.2 that suggested the phenotype difference observed for the facultative versus anaerobic phenotype.

Table 4.2: Genome features of *G. vaginalis* N153, N101, N95, W11 and B482

Genome Name	Genome Size	Gene Count	Scaffold Count	GC %	16S rRNA Count	tRNA Count	COG[1] Count	Pfam[2] Count	KEGG[3] Count
N101	1543396	1234	10	42	1	45	959	1018	407
N95	1522480	1235	7	42	1	45	967	1009	409
B482	1546682	1253	22	42	1	45	971	1038	408
W11	1566657	1258	7	42	1	45	973	1028	407
N153	1541878	1261	7	42	1	45	965	1025	401

[1]COGs - Clusters of Orthologous Groups of proteins database contains proteins from completely sequenced genomes classified on the basis of the orthology concept. Orthologs are proteins that are direct evolutionary counterparts related by vertical descent between organisms.

[2]Pfam - The Pfam database contains collection of protein families, each family represented by *multiple sequence alignments* and *hidden Markov models (HMMs)* containing proteins with similar functional regions, also called as domain, indicating similar functions.

[3]KEGG (Kyoto Encyclopedia of Genes and Genomes) database integrates genomic, chemical and systemic functional information by obtaining gene catalogs from completely sequenced genomes and linking it to higher-level systemic functions of the cell, the organism and the ecosystem.

In order to gain further insight into overall similarities between genomes, pairwise ANIm (Average Nucleotide Identity by mummer algorithm) values of isolates of N153, N101, W11, N95 and B482 were calculated using JSpecies (Table 4.3). ANIm for N153 and N101 was 99%, which was not unexpected since these two anaerobic isolates originated from the same woman. N153 and N101 were 96% similar to N95, W11 and B482. N95, W11 and B482 were also 96% similar to each other. Gene content differences are suggested by the ANIm values, however further in depth analysis was required to decipher the genome content differences responsible for the distinct phenotypes.

Table 4.3: Pairwise whole genome similarities of *G. vaginalis* subgroup B isolates based on ANIm (Average Nucleotide Identity by Mummer).

	N153	**B482**	**N101**	**N95**
W11	96.29	96.48	96.30	96.41
N153		96.80	99.93	96.79
B482			96.80	96.80
N101				96.79

To investigate the similarities at the level of their predicted proteomes for these isolates, COG profiles of N153, N101, N95, W11 and B482 were used to generate a histogram at the functional COG category level (Figure 4.1). Figure 4.1 shows the number of predicted genes assigned to each COG category. COG profiles were very similar among the five subgroup B isolates, with nearly identical distributions observed for the five subgroup B isolates.

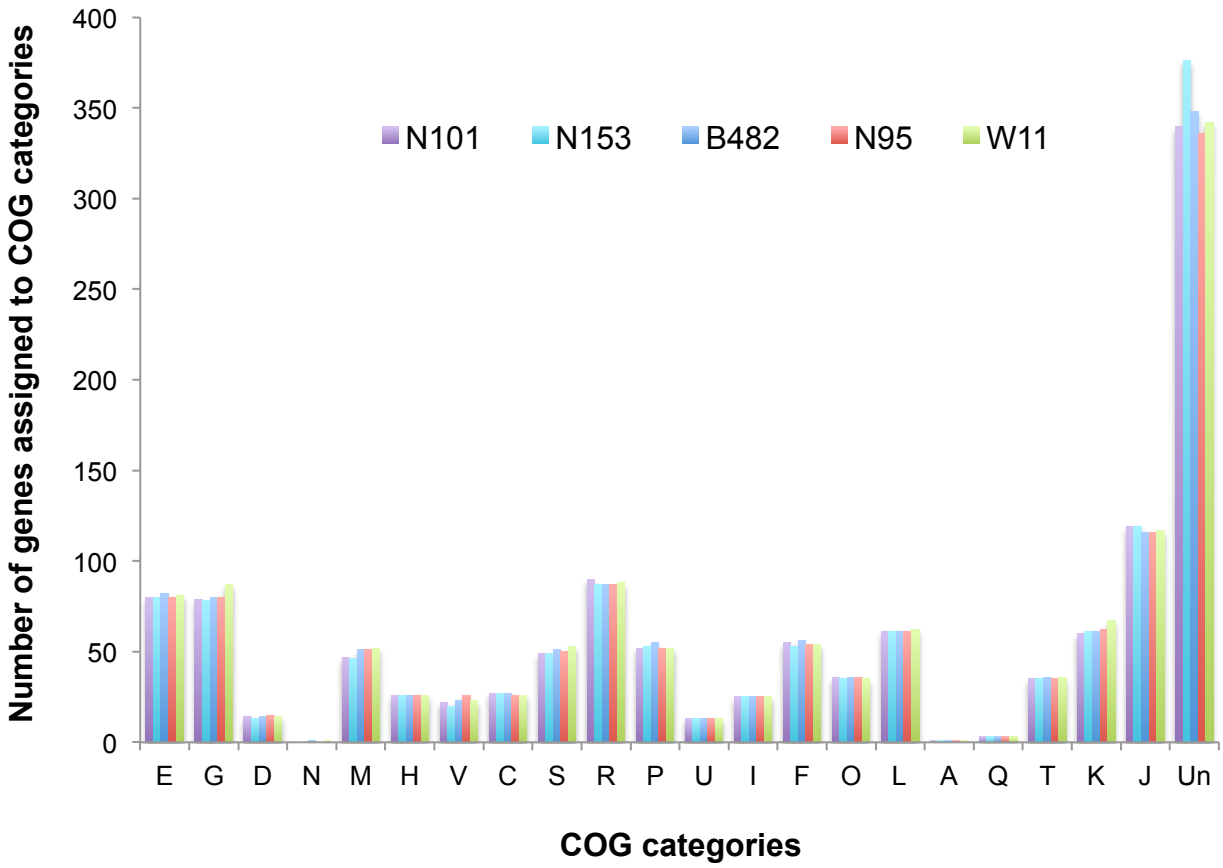


Figure 4.1: Histogram representing genes (number) assigned to COG categories for the facultative and anaerobic genomes.

COG categories are represented using alphabets. *E* - Amino acid transport and metabolism, *G* - Carbohydrate transport and metabolism, *D* - Cell cycle control, cell division, chromosome partitioning, *N* - Cell motility, *M* - Cell wall/membrane/envelope biogenesis, *B* - Chromatin structure and dynamics, *H* - Coenzyme transport and metabolism, *Z* - Cytoskeleton, *V* - Defense mechanisms, *C* - Energy production and conversion, *W* - Extracellular structures, *S* - Function unknown, *R* - General function prediction only, *P* - Inorganic ion transport and metabolism, *U* - Intracellular trafficking, secretion, and vesicular transport , *I* - Lipid transport and metabolism, *Y* - Nuclear structure, *F* - Nucleotide transport and metabolism, *O* - Post-translational modification, protein turnover, and chaperones, *L* - Replication, recombination and repair, *A* - RNA processing and modification, *Q* - Secondary metabolites biosynthesis, transport, and catabolism, *T* - Signal transduction mechanisms, *K* – Transcription and *J* - Translation, ribosomal structure and biogenesis. Un represents ORFs unassigned to any defined COG categories.

For a more detailed comparison of isolates and identification of strain-specific ORFs (Open Reading Frames specific to either facultative strains or to anaerobic strains), *GView* (https://server.gview.ca/) was used to display and navigate bacterial genomes. In order to identify ORFs unique to facultative genomes, the genome of a facultative strain (B482) was used as a seed genome. ORFs of other genomes are compared using tblastn to ORFs of the seed genome and ORFs with percent identity less than 70% over an alignment length of 100 are considered to be unique. The pangenome is then created by iterative addition of unique regions to the initial seed genome such that the resulting pangenome contains ORFs from all genomes. Regions not present in one genome are visible as gaps in the BLAST results for that genome against the pangenome.

From the pangenome analysis, a contiguous 15 kb genome fragment was identified in B482, which was also present in W11 and N95, but was absent in the anaerobic strains N153 and N101. Figure 4.2 illustrates the 15 kb fragment, which contains 13 predicted ORFs that were only present in the facultative strains (N95, W11 and B482). The COG functional categorization of these ORFs indicates that five out of thirteen ORFs are involved with cell wall/membrane/envelope biogenesis. Out of the remaining eight, three are predicted to be involved with carbohydrate metabolism and transport and one ORF is predicted to be a gene involved with lipid transport and metabolism. The COG categories for the remaining ORFs could either not be assigned (category Un) or were predicted for general functionality (category R). This fragment includes a predicted epimerase coding gene, described in the literature to be involved in capsule synthesis (Zuppardo and Siebeling 1998) and also a polysaccharide ABC transporter membrane protein coding gene, which is suggested to be involved in

transporting O-antigen and other sugar moieties such as rhamnose-glucose polysaccharide (synthesized by rhamnan synthesis protein F) to the exterior of the cell, enabling the formation of a capsule like structure. The annotations of these ORFs and their COG affiliations, suggest that this 15 kb fragment is involved in cell wall biosynthesis or repair, and particularly in capsule synthesis.

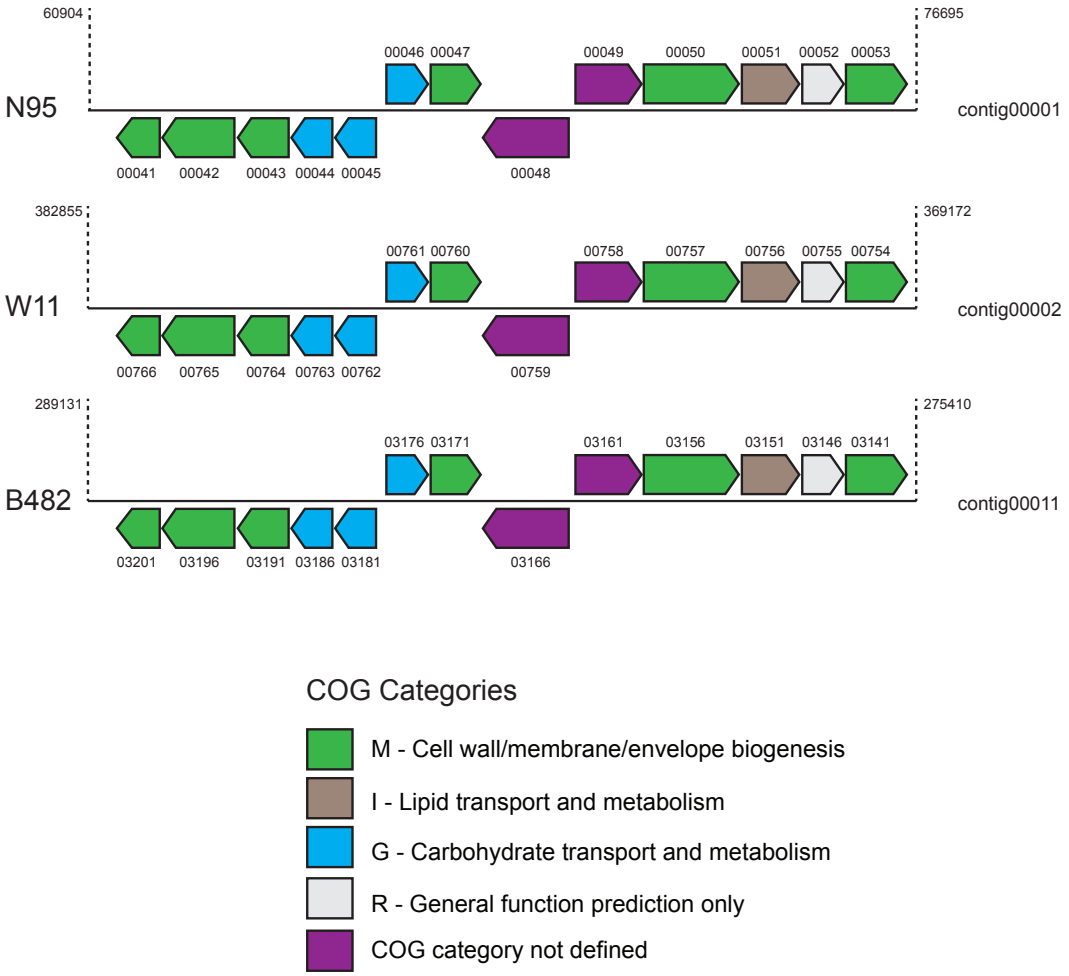


Figure 4.2: A 15 kb fragment present only in facultative strains (N95, W11 and B482) described using COG categories.

Predicted genes from the above mentioned 15 kb fragment and other ORFs that were identified only in the facultative strains, and absent in the anaerobic strains, are listed in Table 4.4. Out of the 21 ORFs listed in the table, 13 ORFs (00041 to 00053 of N95) are located on the 15 kb fragment described above. The remaining 8 ORFs are scattered around the genome. Out of these eight, four (00019, 01005, 00486 and 00480 of N95) could not be assigned to a COG category, and one ORF (00484 in N95) was assigned to COG category R (General function prediction only). As with most of these genes in the unique list, one of the remaining three (00483, COG category M) is predicted to be involved in cell wall/membrane/envelope biogenesis. The remaining two (01023 and 00018 of N95) are interesting candidates to be considered, as they are predicted to be involved in replication, recombination and repair functions. Replication, recombination and repair related genes are important for any cell in protecting the cell against oxidative stress (Henry *et al.* 2012).

Table 4.4. ORFs present in facultative strains of *G. vaginalis* (B482, N95 and Wll) and absent in anaerobic strain (N153 and N101).

COG category	COG	JGI Annotation	ORF # in B482	ORF # in N95 (% identity to ORF in B482)	ORF # in W11 (% identity to ORF in B482)
R General function prediction only	COG3668 Plasmid stabilization system protein	Addiction module toxin, RelE/StbE family protein	04281	00484 (100)	00965 (100)
	COG1216 Predicted glycosyltransferases	Glycosyltransferase family protein	03146	00052 (94.2)	00755 (91.9)
G Carbohydrate transport and metabolism	COG1134 ABC-type polysaccharide/polyol phosphate transport system, ATPase component	ABC-type polysaccharide/polyol phosphate transport system, ATPase component	03186	00044 (100)	00763 (99.6)
M Cell wall/membrane/envelope biogenesis	COG3754 Lipopolysaccharide biosynthesis protein	Lipopolysaccharide biosynthesis protein	03156	00050 (91.5)	00757 (93.8)
	COG1215 Glycosyltransferases, probably involved in cell wall biogenesis	Glycosyltransferases	04286	00483 (99.6)	00964 (99.6)
	COG1209 dTDP-glucose pyrophosphorylase	dTDP-glucose pyrophosphorylase	03201	00041 (100)	00766 (95.0)
	COG1091 dTDP-4-dehydrorhamnose reductase 2E-54	fused dTDP-4-keto-L-rhamnose reductase and dTDP-4-keto-6-deoxyglucose-3,5-epimerase enzyme	03196	00042 (97.5)	00765 (93.8)
	COG1088 dTDP-D-glucose 4,6-dehydratase	dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)	03191	00043 (100)	00764 (95.4)
	COG0463 Glycosyltransferases involved in cell wall biogenesis	Glycosyltransferases involved in cell wall biogenesis	03141	00053 (82.6)	00754 (84.0)
		Glycosyltransferases involved in cell wall biogenesis	03161	00050 (91.5)	00757 (93.8)
		Glycosyltransferases involved in cell wall biogenesis	03171	00047 (96.6)	00760 (100)
L Replication, recombination and repair	COG1195 Recombinational DNA repair ATPase (RecF pathway)	AAA domain	01429	01023 (99.8)	01135 (95.2)
	COG1112 Superfamily, DNA and RNA helicases and helicase subunits	type III restriction protein res subunit	03915	00018 (99.9)	00790 (100)

I Lipid transport and metabolism	COG1835 Predicted acyltransferases	Predicted acyltransferases	03151	00051 (96.9)	00756 (92.9)
G Carbohydrate transport and metabolism | M Cell wall/membrane/envelope biogenesis	COG1682 ABC-type polysaccharide/polyol phosphate export systems, permease component	O-antigen export system permease protein	03181	00045 (100)	00762 (95.8)
	COG0451 Nucleoside-diphosphate-sugar epimerases	Nucleoside-diphosphate-sugar epimerases	03176	00046 (96.8)	00761 (99.3)
-	-	KilA domain protein	03905	00019 (88.0)	00789 (96.9)
-	-	Hypothetical protein	01499	01005 (100)	01161 (100)
-	-	Hypothetical protein	04266	00486 (99.0)	00967 (99.0)
-	-	Hypothetical protein	04296	00480 (99.3)	00962 (99.3)
-	-	Glucosyltransferase GtrII	03166	00048 (98.0)	00759 (98.0)

Genes uniquely present in anaerobic strains, absent in the facultative strains are listed in Table 4.5. Sixteen of these ORFs are contained within a contiguous genome fragment of 21 kb size: ORFs numbered 00039 to 00055 in N153, 00658 to 00672 in N101, with 00660 in N101 corresponding to a fusion of 00041 and 00042 of N153 (Figure 4.3). Five of these ORFs were assigned to COG category M (Cell wall biogenesis). ORFs corresponding to 00047 and 00048 in N153 are predicted to be involved in carbohydrate transport and metabolism (COG category G). ORF 00054 is predicted to be involved in replication, recombination and repair. Another ORF in this fragment (00046 in N153) is predicted to be involved in lipid transport and metabolism. The remaining ORFs in this fragment were assigned to COG category R (General functional predication only) or were unassigned.

Table 4.5. ORFs present in anaerobic strains of *G. vaginalis* (N153 and N101) and absent in facultative strains (N95, W11 and B482)

COG category	COG	Description	ORF # in N153	ORF # in N101 (% identity to ORFs in N153)
V Defense mechanisms	COG1715 Restriction endonuclease	Restriction endonuclease	01110	00481 (100)
R General function prediction only	COG1216 Predicted glycosyltransferases	Predicted glycosyltransferases	00050	00668 (100)
	COG3153 Predicted acetyltransferase	Predicted acetyltransferase	01104	00487 (100)
	COG3668 Plasmid stabilization system protein	Plasmid stabilization system protein	00287	00896 (100)
M Cell wall/membrane/envelope biogenesis	COG0438 Glycosyltransferase	Glycosyltransferase	00049	00667 (100)
			00040	00659 (100)
	COG1088 dTDP-D-glucose 4,6-dehydratase	dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)	00055	00672 (100)
	COG3475 LPS biosynthesis protein	LPS biosynthesis protein	00044	00662 (100)
	COG4750 CTP:phosphocholine cytidylyltransferase involved in choline phosphorylation for cell surface LPS epitopes	CTP:phosphocholine cytidylyltransferase involved in choline phosphorylation for cell surface LPS epitopes	00041	00660 (98.4)
L Replication, recombination and repair	COG0494 NTP pyrophosphohydrolases including oxidative damage repair enzymes	NTP pyrophosphohydrolases including oxidative damage repair enzymes	00054	00671 (100)
	COG3077 DNA-damage-inducible protein J	addiction module antitoxin, RelB/DinJ family	00146	00759 (100)
J Translation, ribosomal structure and biogenesis	COG1670 Acetyltransferases, including N-acetylases of ribosomal proteins	Acetyltransferases, including N-acetylases of ribosomal proteins	00497	01100 (100)
I Lipid transport and metabolism	COG1835 Predicted acyltransferases	Predicted acyltransferases	00046	00664 (100)
G Carbohydrate transport and metabolism	COG1134 ABC-type polysaccharide/polyol phosphate	ABC-type polysaccharide/polyol phosphate transport system,	00048	00666 (100)

	transport system, ATPase component	ATPase component		
	COG1682 ABC-type polysaccharide/polyol phosphate export systems, permease component	ABC-type polysaccharide/polyol phosphate export systems, permease component	00047	00665 (100)
C Energy production and conversion	COG1249 Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide dehydrogenase (E3) component, and related enzymes	Mercuric reductase	00267	00877 (100)
-	-	Hypothetical protein	01261	00412 (100)
-	-	abi-like protein	00554	00408 (100)
-	-	Addiction module toxin, RelE/StbE family	00147	00760 (100)
-	-	BNR repeat-like domain	00939	00029 (99.89)
-	-	Ca protein B-type domain	00556	00406 (100)
-	-	Domain of unknown function (DUF4417)	00496	01099 (100)
-	-	Hypothetical protein	00039	00658 (100)
-	-	Hypothetical protein	00045	00663 (100)
-	-	Hypothetical protein	00052	00670 (99.4)
-	-	Hypothetical protein	00053	00670 (99.4)
-	-	Hypothetical protein	00130	00744 (100)
-	-	Hypothetical protein	00244	00855 (100)
-	-	Hypothetical protein	00495	01098 (100)
-	-	Hypothetical protein	01023	00566 (100)
-	-	Hypothetical protein	01084	00507 (100)
-	-	Hypothetical protein	01102	00489 (100)
-	-	Hypothetical protein	01119	00473 (100)
-	-	Multidrug resistance efflux transporter	00043	00661 (100)
-	-	Phosphotransferase enzyme family	00042	00660 (100)
-	-	Transposase and inactivated derivatives	00268	00878 (100)

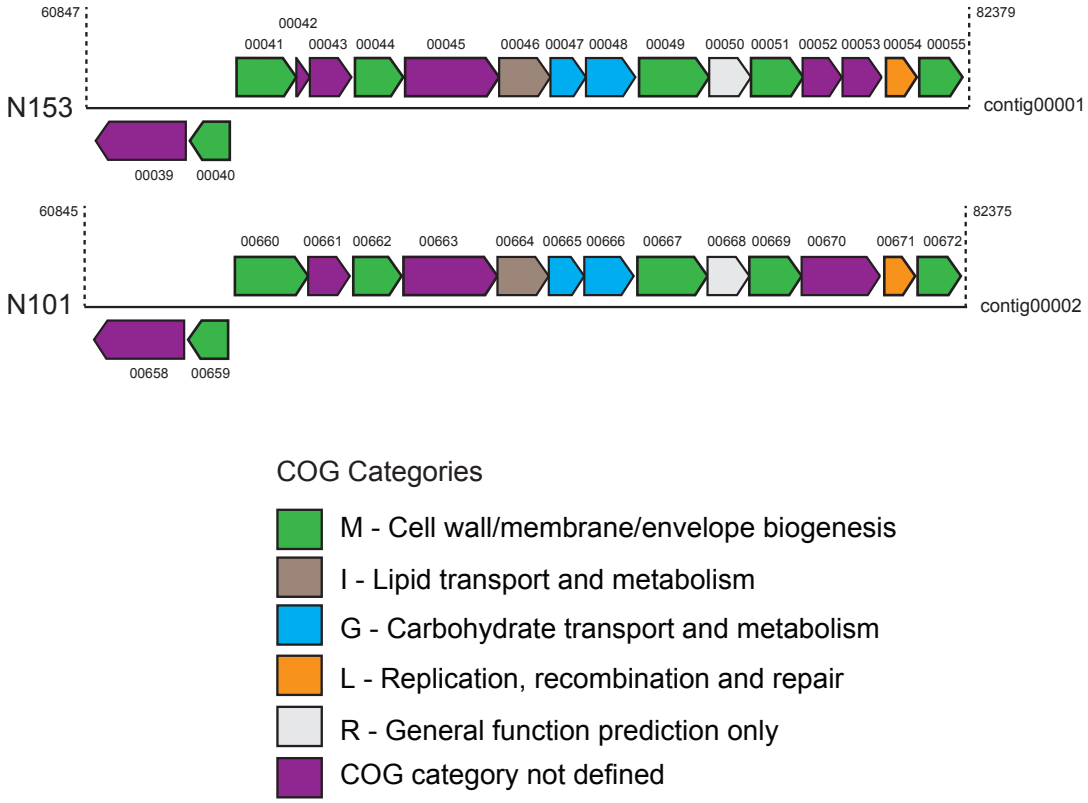


Figure 4.3: A 21 kb fragment present only in anaerobic strains (N153 and N101) described using COG categories.

The majority of the ORFs (14/20) listed in Table 4.5 that are not part of the 21 kb contiguous fragment were not assigned to any COG category. Two of the remaining ones were assigned to COG category R, and four belong to the following COG categories: defense mechanisms (V), replication, recombination and repair (L), translation, ribosomal structure and biogenesis (J) and energy production and conversion (C).

4.5 Discussion

Isolation conditions for the diverse group of organisms designated as *G. vaginalis* vary according to researchers (Ahmed *et al.* 2012; Aroutcheva *et al.* 2001; Lewis *et al.* 2013; Malone *et al.* 1975). The conditions recommended by the American Type Culture Collection (http://www.atcc.org/) for the type strain of this species, ATCC 14018, include incubation in 7% CO_2 atmosphere, however in some recent studies researchers have used anaerobic conditions while trying to isolate this organism (Lewis *et al.* 2013; Schellenberg *et al.* 2012). Malone *et al.* in 1975 isolated obligate anaerobic strains of *G. vaginalis* (Malone *et al.* 1975). Following this observation, researchers tried to isolate more obligate anaerobic strains of *G. vaginalis*, but found that they were not common and were unsuccessful (Bailey *et al.* 1979). The biological relevance of this observation was never understood and no association between obligate anaerobic *G. vaginalis* and phenotyping schemes such as Piot or Benito biotyping has been described.

In their supragenome analysis of 17 *G. vaginalis* isolates, Ahmed *et al.* (2012) did not report any *G. vaginalis* isolates being obligate anaerobes, and all isolates included in that study were cultured on human blood tween (HBT) bilayer agar in 6 % CO_2. One of their isolates, B482, has a *cpn*60 UT sequence identical to N153 and N101, subgroup B isolates that were found in our study to be anaerobic. The genome sizes (1.55 - 1.57 Mb)

and GC content (42%) of the whole genomes of the facultative (N95 and W11) and anaerobic (N153 and N101) described in this study are similar to the subgroup B (Ahmed *et al.* clade 2) isolates described by Ahmed *et al*. in their study of 17 *G. vaginalis* genomes (Ahmed *et al.* 2012). COG categorization of open reading frames for the five genomes was very similar (Figure 4.1) further suggesting that any gene content differences detected between the facultative and anaerobic strains would be related to their different atmospheric requirements.

A pangenome analysis revealed contiguous genome fragments containing open reading frames present uniquely either in facultative strains or anaerobic strains (Figure 4.2 and Figure 4.3). Both contiguous fragments, either unique to anaerobic strains or facultative strains, are speculated to be involved with cell wall biogenesis based on the annotation of ORFs present in these fragments. The genome fragment uniquely present in facultative strains contained genes involved in cell wall biogenesis, carbohydrate and lipid transport and metabolism. This contiguous fragment also contains a predicted epimerase gene and a polysaccharide ABC transporter membrane protein coding gene. Epimerases have been described as critical for capsule synthesis and absence of this function is associated with the disruption of capsule synthesis (Zuppardo and Siebeling 1998). The polysaccharide ABC transporter membrane protein is annotated as having substrate specificity to O-antigen and other sugar moieties such as rhamnose-glucose polysaccharide. O-antigen, also known O-specific polysaccharide, is made of repeating oligosaccharide subunits and can a major constituent of the bacterial outer membrane (Marolda *et al.* 2004). Polysaccharide capsules, composed of repeating oligosaccharides, are found on surfaces of many Gram negative and Gram positive bacteria (Moxon and

Kroll 1990). Capsule producing bacteria are involved in reducing the host proinflammatory immune response and are less subject to phagocytosis by macrophages and dendritic cells than non-encapsulated bacteria (Henry *et al.* 2012).

All organisms have some defense as well as repair mechanisms to protect their proteins, lipids, RNA and DNA from reactive oxygen species (ROS). The protective mechanisms used by organisms vary and are most critical to strict anaerobes. One of the mechanisms described and studied scarcely in the literature is the cell wall and the role of different surface structures such as capsules in protecting the prokaryotes against oxidative stress (Henry *et al.* 2012). In *Porphyromonas gingivalis* (Henry *et al.* 2012) and also in the eukaryote *Cryptococcus neoformans* (yeast) (Zaragoza *et al.* 2008; Zaragoza *et al.* 2009), the presence of a capsule also protects the organism against oxidative stress. In *Cryptococcus neoformans,* the capsule acts as a scavenger of reactive oxidative species enhancing its survival when ingested by phagocytic cells (Zaragoza *et al.* 2008). The 15 kb contiguous genome fragment only present in facultative strains is predicted to be involved in capsule synthesis. Since this fragment is absent in anaerobic strains, the capsule synthesized by this fragment might be involved in protective the facultative strains from oxidative stress and could be responsible for the phenotype of these strains being able to grow at 7% CO_2 atmosphere.

It is interesting to note that another contiguous genome fragment (21 kb fragment in figure 4.3) only present in anaerobic strains is also predicted to be involved with cell wall biogenesis. However, it does not include an epimerase gene and hence is not predicted to be involved in capsule synthesis. The absence of capsule in the obligate

anaerobic strains might be mitigated by the presence of a different cell wall design, allowing its survival in anaerobic conditions.

The cell wall of *G. vaginalis* has always been described with caution, and the organism is referred to as "Gram variable" since it can appear as Gram positive or Gram negative based on growth stage or conditions (Catlin 1992). The cell wall of isolates of *G. vaginalis* including the type strain have been described as having a Gram positive organization, but with multiple laminations mimicking the Gram negative cell wall (Greenwood and Poickett 1980). An exopolysaccharide layer is thought to be involved in helping bacteria adhere to the epithelial cells (Catlin 1992). Harwich *et al.* observed a capsular structure on *G. vaginalis* isolate 5-1 that was not apparent for a closely related strain, AMD (both subgroup A strains). The authors speculated that the differential production of this polysaccharide capsular material might explain the differences observed in biofilm formation by these isolates (Harwich *et al.* 2010). The functionality of polysaccharide capsule in protecting the cells against oxidative stress remains unexplored in this group of organisms.

Horizontal gene transfer via transposons, the CRISPR/Cas system, and phage-mediated gene shuttling through mobile elements have all been reported in *G. vaginalis* (Pleckaityte *et al.* 2012b; Yeoman *et al.* 2010) and some combination of these processes is likely responsible for the differences in genome content we observed in this study. Further investigation into the functions of genes differentially detected in subgroup B isolates in the current study will help explain how they might be involved in protecting cells from oxidation damage, allowing W11, N95 and B482 to grow in concentrations of oxygen as high as 15-20% (found in CO_2 incubators), much higher than the concentration

present in atmosphere generated by anaerobic Gaspaks (<0.1% oxygen). Clarifying the role of *G. vaginalis* in BV and the part it plays in tipping the balance in the vaginal microbial community from *Lactobacillus* dominance to a consortium of anaerobes depends upon resolution of phenotypically distinct subgroups within this diverse taxon and a corresponding shift away from thinking about *G. vaginalis* as a single entity.

CHAPTER 5 - A prospective study of the vaginal microbiota following preterm premature rupture of membranes (PPROM)

Citation

Some of the data included in this chapter have been published in the conference proceedings for the 35th Annual Meeting of the Society for Maternal-Fetal Medicine: The Pregnancy Meeting, San Diego CA, February 2-7, 2015.

Hill JE, Jayaprakash TP, Wagner EC, Albert AYK, van Schalkwyk J, American Journal of Obstetrics and Gynecology, 212(1): S320–S321.

A manuscript with this same author list, and including the data presented in this chapter is in preparation for submission at the time of thesis submission.

Author Contributions

Conceived and designed the experiments: TPJ ECW JvS AYKA JEH DMM. Performed the experiments: TPJ. Analyzed the data: TPJ ECW AYKA JEH. Wrote the paper: TPJ ECW JvS AYKA JEH DMM.

Fifth Chapter Transition

The last three chapters have focused on improving the resolution of the vaginal microbiome specifically with respect to the genus *Gardnerella*. Through the accomplishment of objectives of categorizing *Gardnerella* into four subgroups using the *cpn*60 UT, and also establishing that there are clinical significances associated with these *cpn*60 subgroups, I have demonstrated *cpn*60 can be a very useful tool in providing a comprehensive and high resolution picture of any microbial community.

With that in mind, this chapter has focused on accomplishing the last objective. This chapter is prospective study of vaginal microbiota of women with PPROM using *cpn*60 based microbial profiling to identify any microbiome structure or pattern that can be used to predict maternal or neonatal outcomes.

5.1 Abstract

Preterm premature rupture of membranes (PPROM) is a precursor to up to one third of all preterm births, and is associated with adverse maternal and neonatal outcomes. Delivery following PPROM can occur immediately or following an extended latency, and although abnormal vaginal microbiota has been implicated in triggering PPROM, the role of this microbial community in affecting latency duration is unknown. The objective of this study was to characterize the vaginal microbiota of women following PPROM, and to determine if microbiome composition predicts latency duration and perinatal outcomes. Microbiome profiles based on amplification and pyrosequencing of the *cpn*60 universal target were generated for vaginal swabs from 36 women enrolled in the study, at the time of presentation with PPROM, once per week throughout the latency period, and at delivery. Microbiome profiles were generally dominated by one or two species including *Lactobacillus crispatus, L. iners, Prevotella timonensis, P. bivia, G. vaginalis,* and *Escherichia coli*. Only 13/70 samples examined were dominated by *Lactobacillus* spp., the expected profile for healthy women. Sequences representing *Megasphaera* type 1 and *Prevotella* spp. were detected in all vaginal samples collected throughout the study (n=70). Microbiome profiles at the time of membrane rupture did not cluster by gestational age at PPROM, latency duration, or chorioamnionitis diagnosis. The vaginal microbiota was generally unstable over the latency period, with dramatic shifts in composition between weekly samples, and an overall decrease in *Lactobacillus* abundance. Mollicutes (*Mycoplasma* and/or *Ureaplasma)* were detected by PCR in 81% (29/36) of women, and these women had significantly lower GA at delivery and correspondingly lower birth weight infants than

Mollicutes PCR negative women. The results of this study provide new insight into the vaginal microbiota of women who have experienced PPROM and demonstrate the instability of the microbiome throughout the latency period.

5.2 Background

Preterm premature rupture of the membranes (PPROM) is a precursor to 20-30% of all preterm deliveries (Goldenberg *et al.* 2008; Mercer 2003). The latency period between membrane rupture and delivery is a critical period for determining maternal and neonatal health outcomes, since while longer latencies have been associated with reduced odds of neonatal morbidity, (Frenette *et al.* 2013) there is an associated increase in risk of maternal or fetal infection (Simhan and Canavan 2005; Yudin *et al.* 2009). The interplay between gestational age (GA) at PPROM, latency, and maternal and neonatal morbidity results in a complex set of trade-offs for maternal and neonatal health, and presents profound challenges for management of PPROM during the watchful waiting period following membrane rupture.

Abnormal vaginal microbiota is a recognized risk factor for PPROM (Epstein *et al.* 1998; Hillier *et al.* 1995; Leitich *et al.* 2003). The infectious etiology hypothesis suggests that in women with subclinical or clinical upper genital tract infections, bacteria ascend into the uterus causing an inflammatory response, which can lead to PPROM or preterm labour (Goldenberg *et al.* 2008; Locksmith and Duff 2001). Historically, anaerobes, Gram negative organisms and mycoplasmas have been associated with intrauterine infection (Goldenberg and Culhane 2003; Goldenberg *et al.* 2008; Goldenberg *et al.* 2005), and some specific microbial species have been associated with PPROM (Goldenberg *et al.* 1998). However, investigations have been limited by dependence on culture and isolation of fastidious organisms.

Culture-independent methods based on high throughput DNA sequencing offer an unprecedented opportunity to investigate the microbiological aspects of PPROM. Studies

focusing on microbes in the fetal membranes, placenta and amniotic fluid of women with preterm birth, PPROM and term labour have resulted in the detection of a wider diversity of microbial species than culture dependent studies, as well as the identification of previously unrecognized species (DiGiulio *et al.* 2010; Fortner *et al.* 2014; Jones *et al.* 2009; Wang *et al.* 2013), and results of these studies further support an association between bacterial colonization of fetal membranes and preterm birth. A recent study by Romero *et al.* (2014a) using 16S rRNA gene sequence-based techniques in a longitudinal study to compare vaginal microbial communities of pregnant women who subsequently delivered preterm or term identified no significant differences in microbiome composition or abundance of bacterial taxa between groups. However, differences in prevalence and abundance of organisms such as *Leptotrichia/Sneathia*, BVAB (BV associated bacteria), *Mobiluncus* spp. and *Mycoplasma* spp. in the vaginal microbiomes of women who deliver term or preterm have been reported in other recent studies employing culture-independent methods (Nelson *et al.* 2014; Wen *et al.* 2014).

Information regarding the vaginal microbiota in PPROM is relatively scarce, and the composition and dynamics of the vaginal microbiota through the latency period following PPROM is currently unexplored using culture-independent methods. While there is certainly an urgent need for improved diagnostic tools for the prediction of risk of PPROM, there is an equal need for tools to aid in management of the latency period once PPROM has occurred. The objectives of the current study were to characterize the vaginal microbiota of women presenting with PPROM and throughout their latency, and to determine if vaginal microbial profile at the time of membrane rupture predicts latency duration and outcomes.

5.3 Methods

Study design

This was a prospective study of women with PPROM between 24^{+0} and 33^{+6} weeks gestational age. Patients were recruited at the BC Women's Hospital in Vancouver, Canada and were enrolled in the study if they met inclusion criteria: adequate comprehension of English language to give written informed consent, singleton gestation, no known fetal anomalies or complications distinct from PPROM, and no maternal or neonatal indication for iatrogenic PPROM. Ruptured membranes were confirmed by evidence of vaginal fluid pooling, fluid with an elevated pH demonstrated by a positive nitrazine test, and a positive ferning test.

The study received ethics approval from the University of British Columbia Children's & Women's Research Ethics Board (certificate no. H08-01904).

Three vaginal swabs were collected from each woman at enrolment: for *Chlamydia* and *Gonorrhea* testing, Gram stain assessment using Nugent Score, and for microbiome characterization. Additional samples were collected weekly until delivery, and at delivery (n=2 per sampling time, for Gram stain and microbiome analysis). Vaginal swabs were sent immediately to the BC Children's Hospital and BC Women's Hospital and Health Centre Laboratory for NAAT testing for *C. trachomatis* and *N. gonorrhoeae* (Cook *et al.* 2005; Van Dyck *et al.* 2001) and Nugent score assessment. Swabs for microbiome analysis were stored at -80 °C until processing.

Demographic and clinical characteristics were collected from patient charts. Clinical chorioamnionitis was diagnosed based on maternal pyrexia of >38°C and at least two of the following: fetal tachycardia >160 bpm, maternal tachycardia >120 bpm,

leukocyte count >14000 cells, uterine tenderness and/or foul amniotic fluid. Histopathological chorioamnionitis was diagnosed based on placental pathology (Queiros da Mota *et al.* 2013).

Neonatal outcome data collected from the infant charts included; APGAR score (Pediatrics 2006), MAIN score (Verma *et al.* 1999), birth weight, gender, and admission to the Neonatal Intensive Care Unit (NICU). For neonatal skin microbiome sampling, swabs were taken from behind the earlobe and were processed and analyzed in the same manner as the vaginal swabs.

Microbiome profiling

Total genomic DNA was extracted from vaginal swabs using a magnetic bead-based kit (MagMAX, Life Technologies, Burlington, ON). PCR amplicon libraries were created as described previously (Hill *et al.* 2006a; Schellenberg *et al.* 2009a; Schellenberg *et al.* 2011b) using primers targeting the *cpn*60 universal target, which were modified with multiplexing ID tags (MID). Amplicon sequencing was performed on the 454 GS FLX Titanium and GS Junior sequencing platforms. Emulsion based clonal amplification and pyrosequencing were conducted according to manufacturer's protocols (Roche, Brandford, CT). Total bacterial DNA content per swab sample was estimated based on quantitative PCR targeting the bacterial 16S rRNA gene using previously described methods (Chaban *et al.* 2013).

Initial processing and quality control of raw *cpn*60 sequence data was performed using default on-rig procedures (Roche, Branford, CT). High quality reads were processed as previously described (Chaban *et al.* 2014) using the microbial Profiling Using Metagenomic Assembly (mPUMA) pipeline (http://mpuma.sourceforge.net)

(Links *et al.* 2013) to assemble operational taxonomic units (OTU) and to calculate OTU abundance with Trinity and Bowtie2, respectively. Taxonomic identification of OTU was accomplished through watered-Blast comparison (Schellenberg *et al.* 2009a) of OTU sequences to the cpnDB_nr reference database (downloaded on March 21, 2013 from www.cpndb.ca, (Hill *et al.* 2004)) within mPUMA. Any OTU sequence with less than 55% identity to any reference sequence was presumed to be a non-*cpn*60 sequence and was removed from the analysis. For most analyses, OTU sequences with the same nearest reference database sequence were pooled as nearest neighbor "species" and their abundances combined accordingly.

Detection of Mollicutes

Mollicutes (*Mycoplasma* and *Ureaplasma*) were detected by genus-specific, conventional semi-nested PCR targeting the 16S rRNA gene (van Kuppeveld *et al.* 1992). The primary PCR targeted a 700 bp portion of the 16S rRNA gene using primers GPO-1 and MGSO (van Kuppeveld *et al.* 1992). PCR was performed under the following conditions: 40 cycles of 94°C for 30 s, 64°C for 30 s, and 72°C for 60 s. The secondary PCR used primers My-ins (Yoshida *et al.* 2002) and MGSO, and 2 μl of the primary PCR product as template. Thermocycling parameters included 35 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 60 s.

Ureaplasma species (*U. parvum* and *U. urealyticum*) were detected using a conventional PCR based on the multiple-banded antigen gene with primers UMS-125 and UMA226, which yield products of two different sizes depending on the target species: 403 bp (*Ureaplasma parvum*) or 443 bp (*Ureaplasma urealyticum*) (Teng *et al.* 1995;

Teng *et al.* 1994). Specific detection of *Mycoplasma genitalium* was done using a species-specific conventional PCR targeting the 16S rRNA gene (Jensen *et al.* 2003).

Mollicutes-specific 16S rRNA gene PCR amplicons from 10 samples obtained from 8 women were pooled and pyrosequenced as described for *cpn*60 amplicon libraries.

Phylogenetic analysis

For phylogenetic tree calculation, sequences were aligned with ClustalW (gap opening penalty = 10, gap extension penalty = 0.10) (Thompson *et al.* 1994). Bootstrap analysis was performed using 100 iterations of the sequence alignments. Distance matrices for the multiple sequence alignments were calculated using the F84 algorithm within *dnadist* in the PHYLIP software package (Felsenstein 1996). Phylogenetic trees were constructed using *neighbor* (Saitou and Nei 1987) and *consense*, and visualized using Dendroscope (Huson *et al.* 2007).

Statistical analysis

Demographic and clinical variables of the study cohort: These variables were investigated using non-parametric, descriptive statistics (IBM SPSS, version 21).

Microbiome data: Shannon's diversity index, Chao1 estimated number of species and jackknifed Bray-Curtis dissimilarity were calculated using the Quantitative Insights Into Microbial Ecology (QIIME) package (Caporaso *et al.* 2010). These indices, calculated from data grouped to the nearest neighbour species level, were bootstrapped 100 times at 1000 reads per sample or their sample maximum when fewer than 1000 reads were available. Rarefaction plots of alpha diversity measures (Shannon's diversity index and

Chao1 estimated number of species) were generated to ensure an adequate sampling depth for each sample had been achieved. Average linkage hierarchical clustering was performed based on the proportion of each nearest neighbor species per sample using the vegan package in R (Oksnen *et al.* 2012).

5.4 Results

Maternal clinical data

Table 5.1 summarizes the demographic and clinical characteristics of 36 women enrolled in the study between September 2010 and December 2012. Body mass index (prepregnancy) was negatively correlated to the length of latency ($\rho s = -0.428$, $n = 31$, $p = 0.016$).

Table 5.1. Demographic and clinical data for study participants (n=36).

Age	32.92 ± 4.83 [22-40]
Pre-pregnancy? BMI	23.69 ± 4.56 [17.6-37.3]
Gestational diabetes	5 [16.1%]
Ethnicity	
Asian	9 [25%]
South Asian	1 [2.8%]
Caucasian	20 [55.6%]
Aboriginal	3 [8.3%]
Other	3 [8.3%]
Substance use during pregnancy	
Smoking	7 [19.4%]
Alcohol use	4 [11.1%]
Sexual history	
Marital Status	
Partnered	32 [88.9%]
Single	4 [11.1%]
Sexual partners during pregnancy (n=31)	1
Vaginal intercourse during pregnancy (n=34)	31 [91.2%]
Oral sex received during pregnancy (n=34)	11 [32.4%]
Pregnancy history	
Previous pregnancy	36 [100%]
Number of previous pregnancies (Gravida)	2.7 [1-11]
Live birth	1.08 [0-8]
Preterm birth	0.19 [0-2]
BV and STI history during pregnancy	
Diagnosed with BV (n=34)	2 [5.9%]
Diagnosed with yeast	7 [19.4]
Diagnosed with UTI	6 [7.4%]
Diagnosed with Group B *Streptococcus* (n=34)	2 [5.9%]
Diagnosed with *Chlamydia* (n=34)	0
Diagnosed with *Gonorrhea* (n=34)	0
Antimicrobial use for genital tract infection during pregnancy (n=31)	5 [16.1]
Other	
Feminine wipes or deodorant products (n=34)	5 [14.7%]

BMI; body mass index
BV; bacterial vaginosis
STI; sexually transmitted infection
Continuous variables are reported as means ± standard deviation [range]
Categorical variables are reported as N [%]

Gestational age (GA) at PPROM, latency duration, and delivery mode are summarized in Figure 5.1 and Table 5.2. The mean GA at PPROM was 29 weeks (median = 29, range 24-34), and the mean GA at delivery was 32 weeks (median = 32, range 25-39). The mean latency period was 18 days (median = 11, range 1-94). As expected, latency (days) was negatively correlated ($\rho s = -0.390$, $n = 36$, $p = 0.019$) with GA at PPROM since the number of potential latency days decreases with increasing GA at PPROM. Since latency duration was a variable of particular interest in this study, we considered both absolute latency duration (<48 hours, between 48 hours and ≤ 7 days, and > 7 days following membrane rupture), and latency duration as a proportion of maximum possible latency. According to the guidelines of the American College of Obstetrics (2007) and Gynecologists and Society of Obstetrics and Gynecologists of Canada (Yudin *et al.* 2009), conservative/expectant management of PPROM is recommended for gestational ages 24 to 33 weeks and induction of labor is warranted after 34 weeks. Thus, the composite variable, "proportional latency", was calculated using the following equation:

Equation 5.1. Proportional latency with 34 weeks as maximum gestational age at delivery

$$\text{Proportional latency} = \frac{GA_{\text{delivery}} - GA_{\text{PPROM}}}{34 - GA_{\text{PPROM}}}$$

Table 5.2. Summary of PPROM to delivery timelines.

	No. of women (/36)
Gestational age at PPROM	
Extremely preterm (<28 weeks)	14
Very preterm (28 to <32 weeks)	17
Moderate preterm (32 to <34 weeks)	5
Absolute latency	
<48 hours	3
Between 48 hours and < 7days	11
> 7 days	22
Proportional latency[1]	
0 – 0.33	15
0.34 – 0.66	10
0.67 – 1.0	11
Mode of delivery	
Vaginal delivery	24
Cesarean delivery	12

[1]Proportional latency calculation is described in the text.

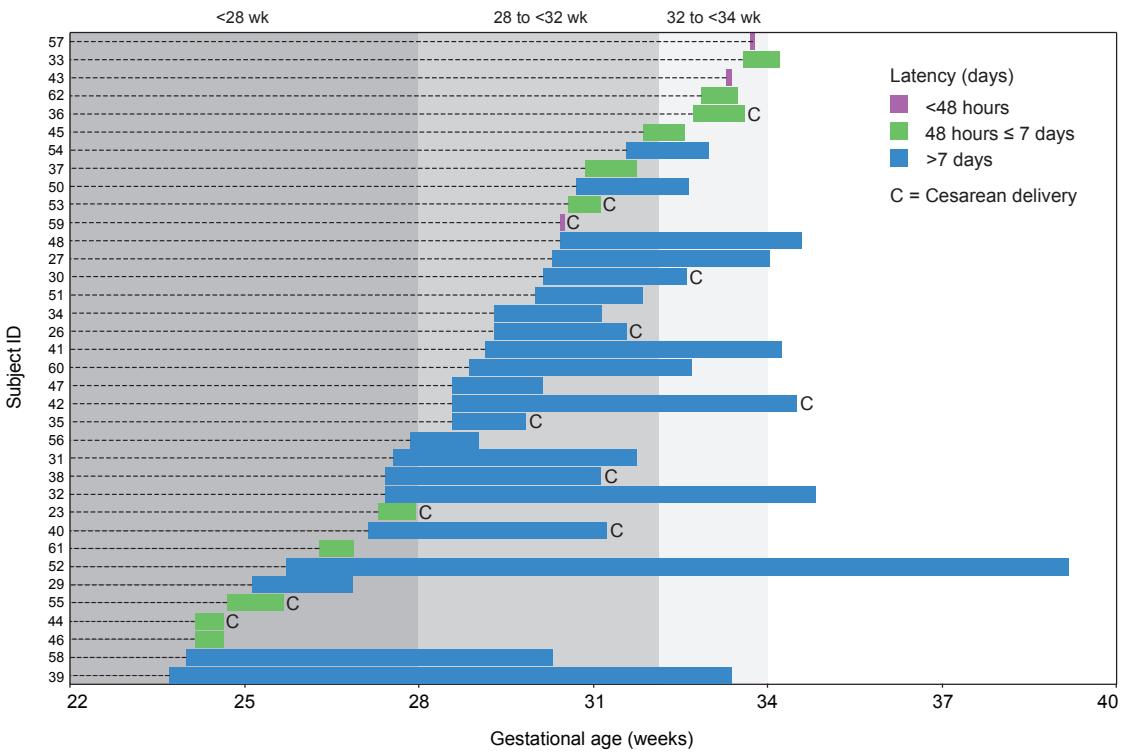


Figure 5.1. GA at PPROM, latency duration and delivery mode for 36 women enrolled in the study. GA ranges corresponding to extremely (<28 weeks), very (28 to <32 weeks) or moderately (32 to <34 weeks) preterm are indicated by grey shading.

Proportional latency, similar to absolute latency, was negatively correlated with BMI (ρs = -0.511, n = 31, p = 0.003). No other significant relationships were identified between proportional or absolute latency and any of the characteristics reported in Table 5.1.

Clinical chorioamnionitis was detected in 35.5% of women (n=36, positive = 11, negative = 20 and data missing = 5). Histopathological chorioamnionitis was detected in 67.7% of women (n=36, histo positive = 21, negative = 10 and data missing = 5).

Samples for microbiome analysis

Over the course of study, 82 samples were obtained from 36 women and their infants at BC Women's Hospital & Health Centre. Vaginal samples collected within 24 hours following PPROM and prior to administration of antibiotics were categorized as T_0 samples (time zero, n = 24). Weekly samples were collected following PPROM until delivery, at which time a delivery sample was collected if possible. The average number of samples per woman was 2 (range, 1 to 7).

Nugent scoring

Nugent scoring of 23 T_0 samples indicated that none of the samples were consistent with BV (Nugent score $\geq$ 7 is consistent with BV). Most (73.91%, 17/23) had scores < 4, while 2/23 (8.7%) had intermediate scores (Nugent score 4-6), and 4/23 (17.4%) could not be scored due to the absence of cells on the Gram stain slides. A total of 43 weekly samples were assessed using the Nugent scoring method, and only one of these had a score consistent with BV. Of the remaining samples, 13/43 (30.2%) had scores < 4, 25/43 (58.1%) had intermediate scores, and 4/43 (9.3%) did not have any cells

on the Gram stain slides. Out of the seven delivery samples, five had intermediate scores and the remaining two had scores of < 4.

Total 16S rRNA and *cpn*60 profiles

Total bacterial content per sample was estimated using a quantitative PCR targeting the bacterial 16S rRNA gene. The overall average value of $\log_{10}$ copy numbers of 16S rRNA gene for vaginal samples was 6.7. Log_{10} copy numbers of 16S rRNA gene per swab for T_0 (n = 24) samples and for weekly samples (n=31) ranged between 4.4 and 8.6, and 4.3 and 8.5, respectively. For delivery samples, $\log_{10}$ copies of 16S rRNA gene ranged between 5.0 and 8.7 per swab. The four T_0 samples that could not be Nugent scored due to the absence of cells had $\log_{10}$ copy numbers of 16S rRNA gene between 4.8 and 6.0. Estimates of total 16S rRNA gene copies in swabs from T_0 samples (mean =6.65 $\log_{10}$ copies per swab) were significantly different (independent T test, t = 5.944, df = 103) from values obtained from 100 clinically healthy pregnant women at low risk for preterm birth (mean = 8.14 $\log_{10}$ copies per swab), which were estimated using the same sample preparation and quantitative PCR method (unpublished data).

For the 70 vaginal swab samples available for *cpn*60-based microbiome profiling, an average of 12,675 *cpn*60 sequence reads were obtained per sample (range 36 - 99, 358; median 6,296). One sample (week 2 sample from woman 60) that yielded only 36 sequence reads was excluded from statistical analyses. Rarefaction analysis of alpha diversity measures (Shannon's diversity index and Chao1 estimated number of species) indicated that all other samples had received adequate coverage for inclusion in further analysis. Shannon's diversity index and Chao1 diversity of T_0 samples were not

significantly different between proportional latency and absolute latency groups as defined in Table 5.2 (Kruskal-Wallis, $p > 0.05$).

A total of 718 OTUs were detected in 82 samples (70 vaginal samples and 12 neonate skin microbiome samples). Of the 670 OTUs detected in the vaginal samples, 9 OTUs accounted for more than 50% of sequence reads obtained. When OTUs were classified according to their nearest neighbour in cpnDB_nr, 20 nearest neighbour taxa (each accounting for at least 1% of the total sequences obtained in the study) accounted for 77% of the data (Table 5.3). The most abundant OTU accounted for 12.4% of the sequence reads, was 97.1% identical to *Prevotella timonensis* JCM 15640 (Table 5.3) and was detected in 34/36 women (64/70 samples).

Table 5.3. Prevalence of nearest neighbour species accounting for at least 1% of the total sequence reads in the study.

Nearest neighbour	No. of OTU	% identity	No. of positive subjects (/36)	No. of positive samples (/70)	Proportion of data (%)
Megasphaera sp. UPII 199-6	1	99.6	36	70	3.0
Lactobacillus iners DSM 13335	2	93.9-99.3	36	60	8.8
Lactobacillus crispatus CECT4840	2	95.3-99.8	35	62	5.8
Prevotella timonensis JCM 15640	1	97.1	34	64	12.4
Gardnerella vaginalis 409-05 (Group A)	2	88.8-99.5	31	51	11.3
Escherichia coli KTE143	1	99.1	31	50	1.8
Lactobacillus jensenii ATCC 25258	2	98.9-99.2	30	43	4.2
Corynebacterium accolens ATCC 49725	4	91.2	29	59	2.6
Gardnerella vaginalis ATCC 14018 (Group C)	1	99.5	29	48	3.0
Streptococcus pseudopneumoniae ATCC BAA-960	3	94.2-94.4	29	45	2.6
Alloscardovia omnicolens LMG 23791	1	97.3	27	46	3.6
BVAB2	2	99.6-99.8	27	38	1.38
Prevotella bivia JCM 6331	1	99.1	25	50	2.6
Peptoniphilus harei ACS-146-V-Sch2b	4	73.6-98.6	25	45	1.3
Corynebacterium striatum ATCC 6940	1	94	22	47	1.4
Bifidobacterium infantis 6w-50	1	99.3	22	35	2.5
Actinomyces sp.	4	74.1-100	21	36	2.7
Streptococcus oralis ATCC 35037	2	88.7-93.8	17	25	1.0
Prevotella veroralis ATCC 33779	1	98.2	15	27	3.1
Streptococcus agalactiae ATCC 13813	1	98.7	12	22	1.8

Average linkage hierarchical clustering of T_0 microbiome profiles (n = 24) was conducted based on nearest neighbour taxa constituting at least 10% of the microbiome in at least one sample (Figure 5.2). The microbiome profiles clustered into profile types, each dominated by one or two species including *Lactobacillus crispatus, Lactobacillus iners, Prevotella timonensis, G. vaginalis, Corynebacterium* sp. and *Escherichia coli* (Figure 5.2). Half (12/24) of T_0 samples were dominated (≥50% of the microbiome) by one or more species of *Lactobacillus*. In eight of these cases, a single *Lactobacillus* sp. comprised at least 50% of the microbiome: *L. crispatus* (n=4), *L. iners* (n=3) and *L. jensenii* (n=1).

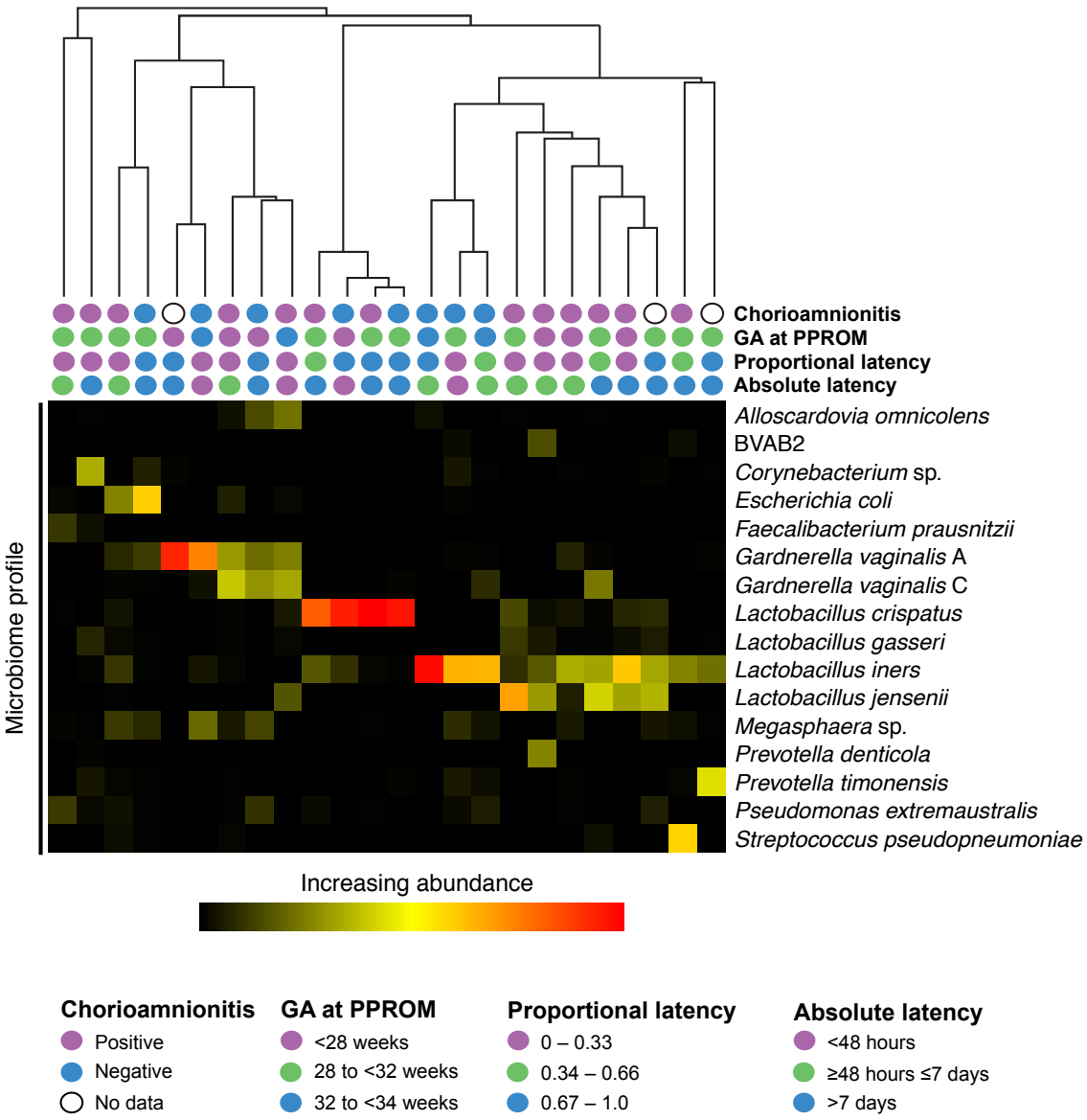


Figure 5.2. Average linkage hierarchical clustering of *cpn*60 based microbiome profiles of T_0 vaginal samples from 24 women.
Nearest neighbour species representing at least 10% of the microbiome of at least one woman are included.

Microbiome profiles did not cluster based on GA at PPROM, latency (absolute latency category or proportional latency), or chorioamnionitis (clinical or histopathological) (Figure 5.2).

Core microbiome and highly prevalent organisms

An OTU sequence (comp15864_c0_seq1) with 99.6% identity to *Megasphaera* sp. UPII 199-6, a type I *Megasphaera* (cpndb ID: b19134, GenBank accession AFIJ01000035), was detected in all 70 vaginal samples and all 12 infant skin samples (comprising 0.0076% to 22.6% of individual microbiome profiles). OTU sequences with best matches to any *Megasphaera* sp. were combined with reference sequences to construct the phylogenetic tree shown in Figure 5.3. Four additional OTU sequences clustered with *Megasphaera micronuciformis* F0359*, Megasphaera elsdenii,* and *Megasphaera* sp. BV3C16-1. One OTU sequence (comp15950_c5_seq1) that was detected in three women had less than 80% identity to any reference *cpn*60 sequence and did not cluster with any known *Megasphaera* species in the phylogenetic analysis.

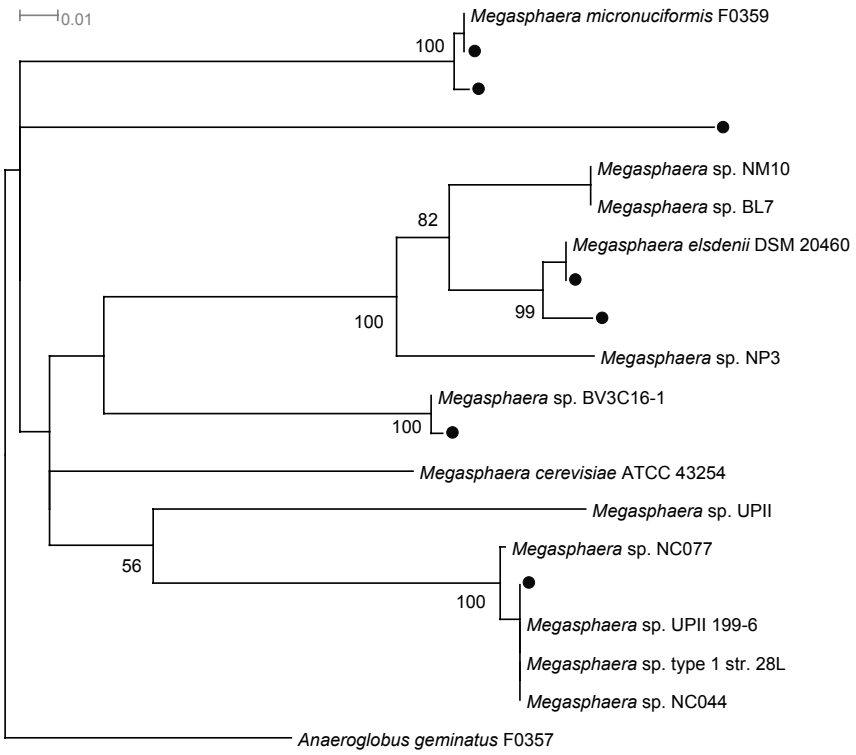


Figure 5.3. Phylogenetic tree of *Megasphaera*-like study sequences (represented by circles, n=6) combined with *Megasphaera* reference sequences.
The tree is based on a 315 bp alignment. Bootstrap values (>50%) are indicated at nodes.

OTU sequences identical to *L. iners* DSM 13335 (cpndb ID: b15282, GenBank Accession: ACLN01000008) were detected in all T_0 samples (n=24). In three of these samples, *L. iners* was the dominant species, accounting for >50% of the total reads.

Prevotella spp. were detected in all vaginal samples, and were also dominant (accounting for ≥50% of reads) in 1/24 T_0 samples and 10/46 weekly samples. Fourteen different species of *Prevotella* were detected in the T_0 samples (with identification based on >95% sequence identity to the reference sequence), along with 15 other *Prevotella*-like OTU sequences with weaker (80-95%) similarity to reference sequences. The most prevalent species were *P. timonensis* and *P. bivia*, which were detected in 92% and 42% of T_0 samples, respectively. Taken together, OTU sequences with affiliation to *Prevotella* were detected in all T_0 samples. In addition to the 29 OTU sequences that were detected in T_0 samples, seven additional *Prevotella*-like OTU sequences were detected in weekly or delivery vaginal samples. *Prevotella*-like OTU sequences were combined with *Prevotella* reference sequences for phylogenetic analysis (Figure 5.4). In addition to the OTU sequences that were closely affiliated to known species, 15 OTU sequences did not cluster with any of the reference sequences, suggesting possibly novel taxa.

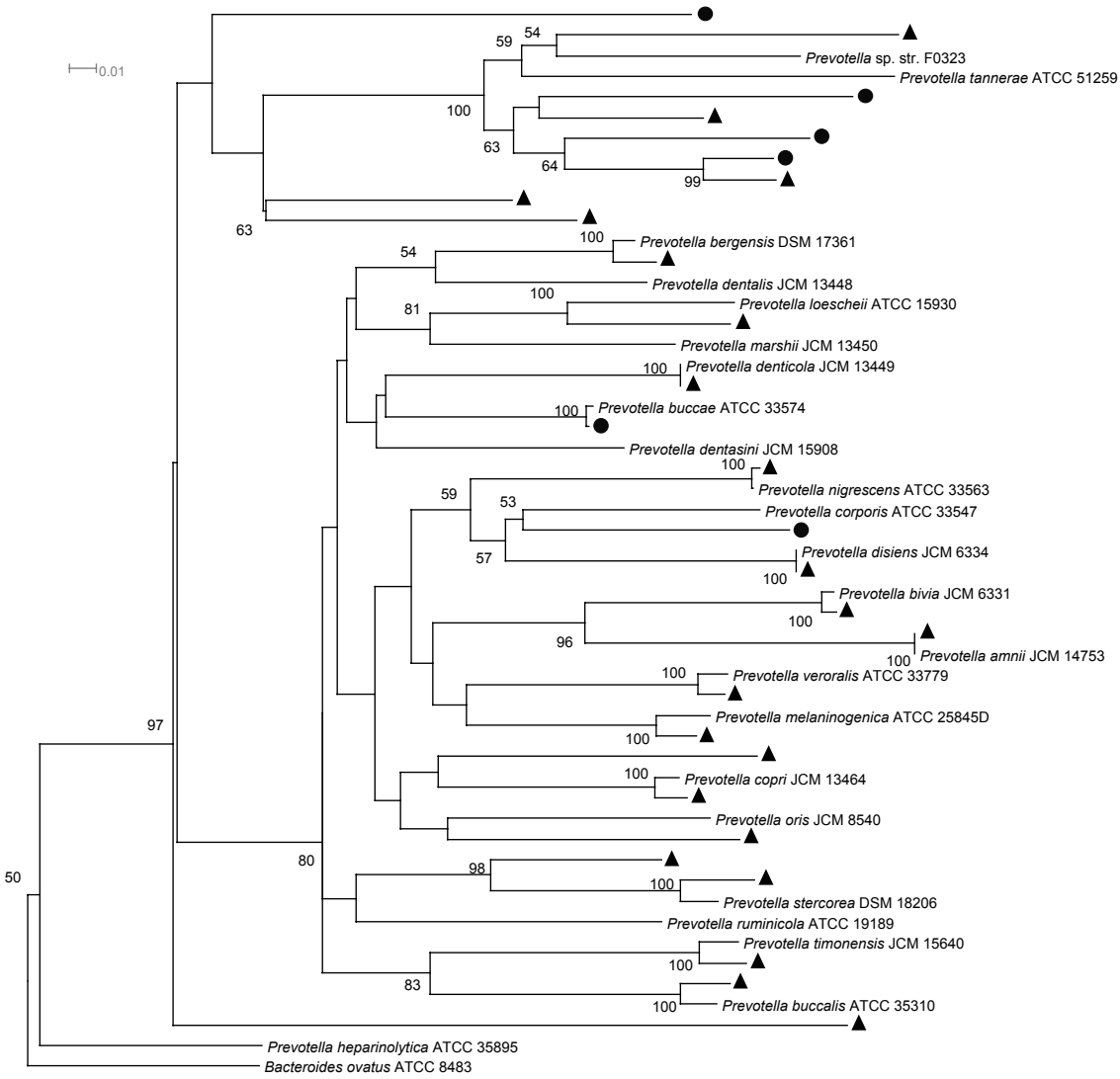


Figure 5.4. Phylogenetic tree of *Prevotella*-like study sequences and *Prevotella* reference sequences. Study sequences detected in T_0 samples are indicated by triangles. Sequences represented by circles were detected only in weekly or delivery vaginal samples. The tree is based on a 300 bp alignment. Only 28 of the 36 identified *Prevotella*-like OTU could be included in the tree since the remaining sequences did not provide sufficient overlap to be included in the alignment. Bootstrap values (>50%) are indicated at node.

Mollicutes

Mollicutes (*Mycoplasma* and/or *Ureaplasma)* were detected by PCR in 81% (29/36) of women (Table 5.3). A woman was considered positive for Mollicutes if any one or more of her samples was found positive by PCR.

Ureaplasma spp. were detected in 53% (19/36) of women, with *Ureaplasma parvum* detected in 39% (14/36) and *Ureaplasma urealyticum* detected in 17% (5/36) of women. Both species were detected together in only one sample. When only T_0 samples were considered, 24/36 (66.7%) were PCR positive for Mollicutes. Non-parametric statistics were used to determine the relationships of proportional latency, absolute latency and GA at birth with Mollicutes detection. Both proportional latency and absolute latency were significantly shorter in women whose T_0 samples were PCR positive for Mollicutes compared to women who were Mollicutes negative (Figure 5.5A, 5.5B), and women whose T_0 samples were positive for Mollicutes had significantly lower GA at delivery (Figure 5.5C), and delivered correspondingly lower birth weight infants (Figure 5.5D).

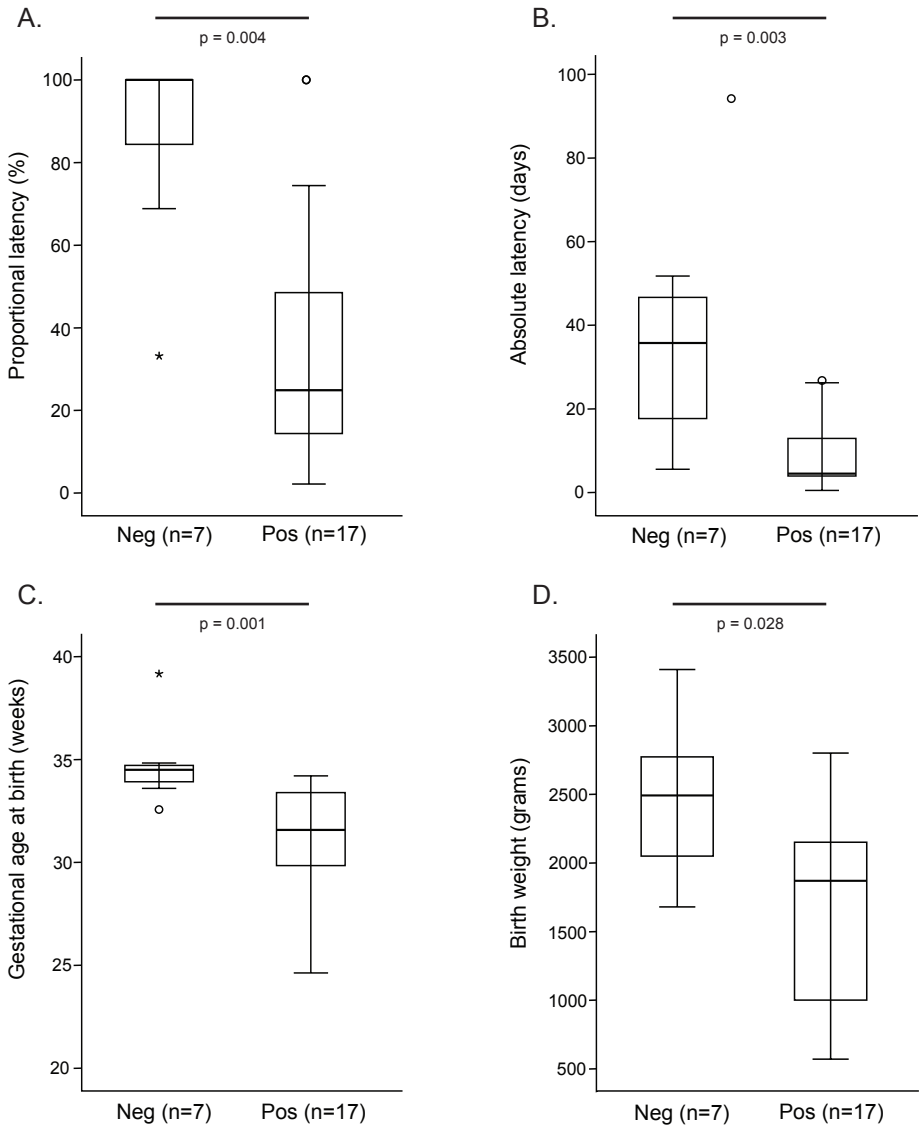


Figure 5.5. Proportional latency (A), absolute latency (B), GA at birth (C) and infant birthweight (D) for women with Mollicutes PCR positive or negative T_0 vaginal samples.
Distributions were compared using a Mann-Whitney U test with $P < 0.05$ considered significant.

To investigate the species of Mollicutes in these samples more thoroughly and to determine if any novel or unexpected species were present, Mollicutes-specific 16S rRNA secondary PCR products from 10 samples (representing 8 women) were pooled and sequenced, resulting in 28,278 high quality sequence reads. All sequence reads were classified (93.7 to 99.1% identical to the reference sequence) as *Mycoplasma hominis*, *Ureaplasma parvum* or *Ureaplasma urealyticum*. None of the reads were classified as *M. genitalium*, and all samples subsequently tested negative by *Mycoplasma genitalium* species-specific PCR, suggesting that *M. genitalium* was not present at detectable levels in these women. Although the Mollicutes-specific primer set is reported to amplify some Firmicutes (Chaban *et al.* 2014), only four reads were identified as non-Mollicutes (97-98% identical to *L. iners*).

Instability of the vaginal microbiota over the latency period

Figure 6 illustrates changes in vaginal microbiome profiles in the weeks following PPROM. All women received one or more broad-spectrum antibiotics (Ampicillin, Erythromycin, Amoxicillin, Metronidazole, Clindamycin, Cefazolin, Nitrofurantoin, and/or Penicillin G). The number of different antibiotics administered to any individual woman during the course of latency ranged from 1 to 7. In all women for whom multiple samples were available, the apparent composition of the vaginal microbiota changed dramatically over the latency period (Figure 5.6). In 4/5 cases where *Lactobacillus* was dominant in the T_0 microbiome, its proportional abundance had decreased substantially by the subsequent sample time point (women 42, 34, 60, 48). The exception to this pattern was woman 51, whose microbiome was dominated by *L. jensenii* and *L. iners* at T_0 and at week 1. Eight of the 19 women for whom multiple samples were available, had

one or more samples dominated (≥50% of the microbiome) by *Prevotella* spp. (Figure 5.6). Of the remaining 11 women, five of them had *Prevotella* spp. as the proportionally most abundant taxon in at least one sample.

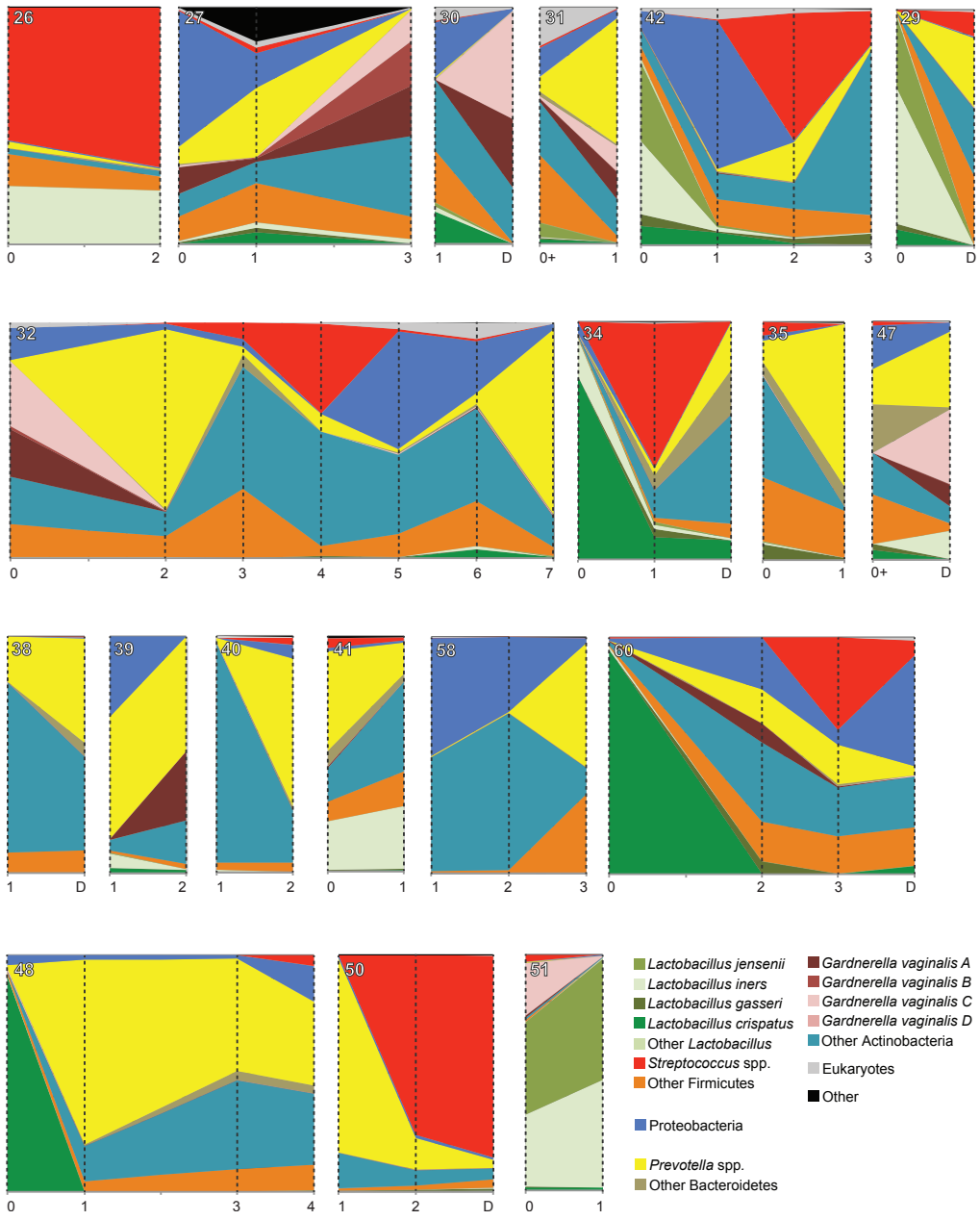


Figure 5.6. Vaginal microbiome profiles over the post-PPROM latency period.
Data is presented as proportion of the total sequence reads obtained for each sample, with the height of the ordinate corresponding to 100%. Sampling times are indicated with vertical broken lines, and collection time (weeks) for each sample is indicated on the abscissa. Samples indicated as collected at "0+" were collected within 3 days following PPROM, but were post-antibiotic treatment. Sample identification numbers appear in the upper left corner of each panel. The legend includes nearest neighbour species that account for at least 10% of the sequence reads in at least one sample.

In addition to the graphical analysis in Figure 5.6, instability of the vaginal microbial community was described quantitatively using the Bray-Curtis dissimilarity index. For women with PPROM, ecological distance between samples within individuals (mean = 0.79, median = 0.86, range = 0.16 - 0.99) was significantly lower (ANOVA, $F_{1, 1828}$ =21.04, p =0.002) than distances between samples from different women (mean = 0.88, median = 0.93, range = 0.10 - 1.00) (Figure 5.7). However, when the PPROM cohort was compared to a cohort of healthy, non-pregnant women who were sampled weekly throughout a menstrual cycle (Chaban *et al.* 2014), the distance between samples within individuals in the PPROM cohort was significantly greater (ANOVA, $F_{1, 112}$ = 95.78, $p < 0.0001$) than the average distance for samples within individuals in the healthy, non-pregnant cohort (mean = 0.29, median = 0.13, range = 0 - 0.98), indicating the greater instability of the PPROM associated microbiomes. These analyses should be considered exploratory as they do not take into account repeated sampling of individual women.

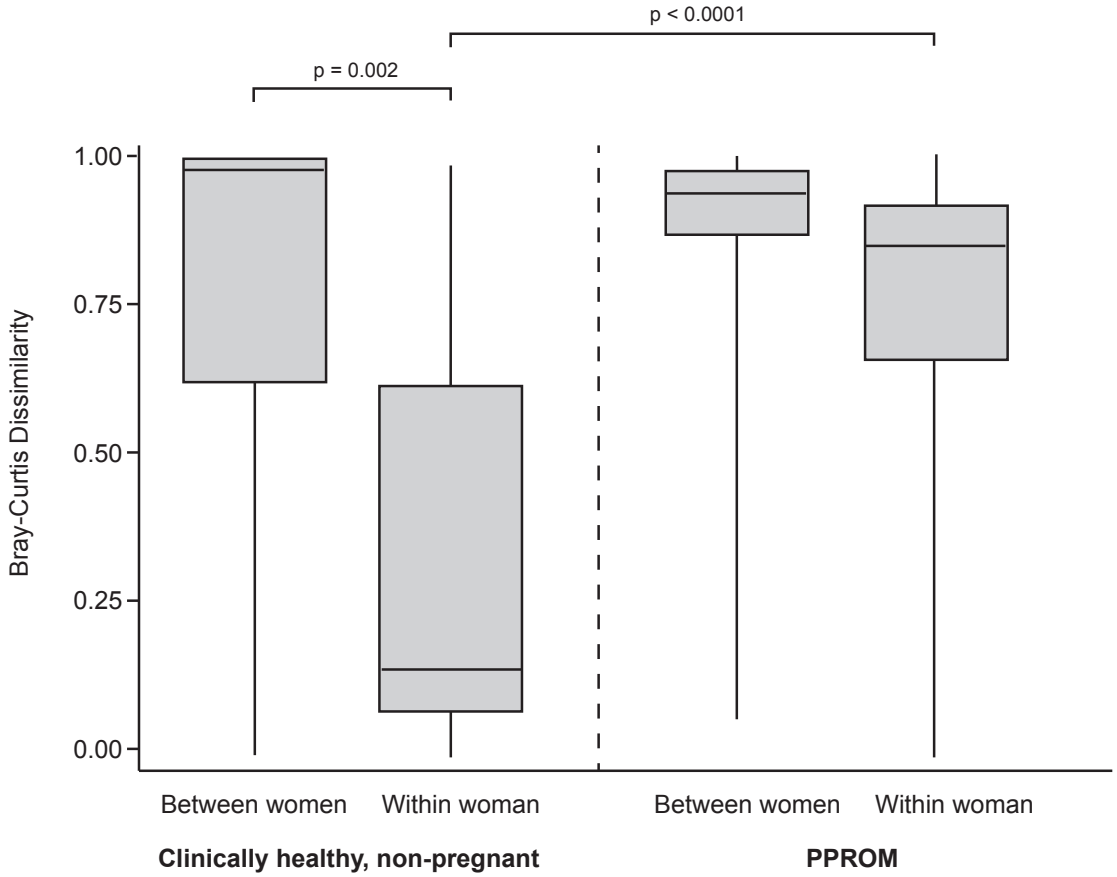


Figure 5.7. Box and whisker plot representing Bray-Curtis distances.
Bray-Curtis distances between vaginal microbiome profiles of women with PPROM, and clinically healthy, non-pregnant reproductive aged women (Chaban *et al.* 2014).

Neonatal outcomes

Table 5.5 summarizes neonatal outcomes. Four babies (4/36, 11.1%) were small for gestational age (<10th percentile). Infants born to women whose T_0 samples were positive for Mollicutes had significantly lower GA at delivery and correspondingly lower birth weight than those born to women with Mollicutes negative T_0 samples (Figure 5C and 5D). Apgar score at 1 minute was positively correlated to gestational age at delivery (ρs = 0.631, n = 35, p = 0.0001), as was Apgar score at 5 minutes (ρs = 0.592, n = 35, p = 0.0001). Since Apgar scores are not usually very evenly distributed, it was categorized into two: <7 and 7-10 and it was found that gestational age at delivery for <7 Apgar score at 1 min was significantly different from the gestational age at delivery for 7-10 Apgar score category.

Table 5.5. Neonatal outcomes.

Gender	
Male	25
Female	11
Birth weight (grams)	1812.9 [570 – 3410]
Extremely low birth weight (<1000 g)	6 [16.7%]
Very low birth weight (<1500 g)	4 [11.1%]
Low birth weight (<2500 g)	20 [55.6%]
Normal	6 [16.7%]
Small for gestational age (n=35)	
Moderate (<10th percentile)	2 [5.7%]
Severe (<3rd percentile)	2 [5.7%]
Bronchopulmonary dysplasia (n=32)	
None	23 [71.8%]
Mild	4 [12.5%]
Severe	3 [9.4%]
Apgar 1 minute score	5.9 [1- 9]
Apgar 5 minute score	7.8 [3-10]
MAIN score	563.2 [0 - 1878]
NICU admission (n=34)	23 [68]
Days of stay in NICU	24.57 [0.81 – 137.74]

Continuous variables are reported as means ± 95% CI [range]
Categorical variables are reported as N [%]

In 11/12 infant skin microbiome profiles, Proteobacteria was the dominant phylum. In the remaining sample, an OTU with 94.2% identity to *Streptococcus pseudopneumoniae* dominated the microbiome. The most prevalent and abundant OTU detected in the skin microbiome of infants included those with similarity to *Escherichia coli* (99.1% identity, 11/12), *Pseudomonas extremaustralis* (95.7% identity, 12/12), *Rhodococcus erythropolis* (100% identity, 12/12), *Streptococcus pseudopneumoniae* (94.2% identity, 8/12), *Acidovorax* sp. (94.1% identity, 12/12), *G. vaginalis* group C (99.1% identity, 5/12), *Prevotella bivia* (99.1 % identity, 8/12), and *Prevotella timonensis* (97.1% identity, 12/12). Sequences corresponding to vaginal microbiome associated organisms *Megasphaera* sp. Type 1 (99.6% identity, 12/12), *Lactobacillus crispatus* (99.8% identity, 12/12) and *L. iners* (99.3% identity, 11/12) were also highly prevalent, but abundance was low (<2% of any individual infant skin microbiome). Mollicutes were detected in 33% (4/12) of the infant skin samples, with two samples PCR positive for *Ureaplasma parvum* and one PCR positive for *Ureaplasma urealyticum*.

Comparisons of vaginal microbiome profiles and infant skin microbiome profiles for 12 mother-infant pairs are shown in Figure 5.8. In each case, the vaginal sample obtained closest to delivery was chosen for comparison. For 7 out of 12 mother-infant pairs, which includes two cases where delivery was by C-section, the Bray Curtis distance between samples was less than 0.79 (the mean Bray Curtis distance between vaginal samples collected from individual women).

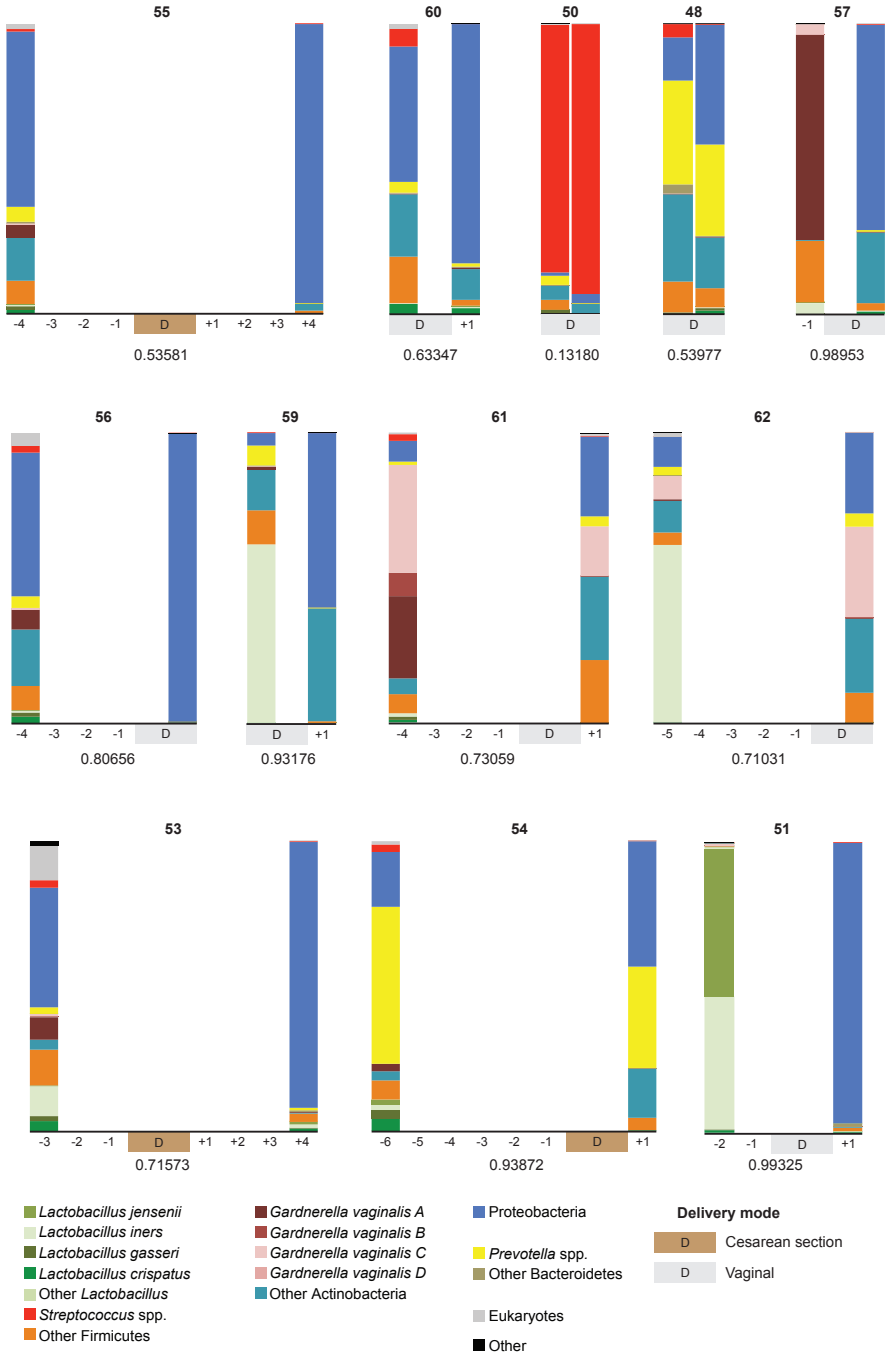


Figure 5.8. Paired maternal vaginal microbiome profiles and infant skin microbiome profiles.
Data is presented as proportion of the total sequence reads obtained for each sample, with the height of the ordinate corresponding to 100%. Days pre-delivery and post delivery are indicated on the abscissa. The day of delivery is indicated with "D". Maternal sample identification numbers appear above each panel. Bray-curtis distances between vaginal sample profile and corresponding infant skin profile are indicated below each panel. The legend includes nearest neighbour species that account for at least 10% of the sequence reads in at least one sample.

All but one of the infants born to the women enrolled in the study survived. One infant, delivered vaginally at 25 weeks after a three-day latency period (birth weight 570 g), died of late onset sepsis at 46 days of life. Blood culture of the infant tested positive for *E. coli* at 12 days of life. *E. coli* was present in the vaginal microbiome of mother (T_0 vaginal sample, detected at level of 0.02% of the total microbiome). The microbiome of the mother had *L. iners* as the most dominant species (32% of reads), followed by *G. vaginalis*. The vaginal T_0 sample tested positive for Mollicutes, and was found to be positive for *Ureaplasma urealyticum*. No infant skin microbiome sample was collected.

5.5 Discussion

Imbalances in the complex microbial community residing in the vagina have been associated with late miscarriages (gestation ages between 16 - 24 weeks) (Hay *et al.* 1994), preterm delivery (Foxman *et al.* 2013; Hillier *et al.* 1995; Holst *et al.* 1994; Martius *et al.* 1988) and PPROM (Gravett *et al.* 1986; Joesoef *et al.* 1995). Disturbed vaginal microbiota is also correlated to neonatal morbidity and adverse maternal outcomes such as clinical chorioamnionitis, postpartum endometritis and wound infections (Gravett *et al.* 1986; Jacobsson *et al.* 2002; Mass *et al.* 1999; Zhang *et al.* 2002). However, little is known about characteristics of the vaginal microbiome during the latency period following PPROM, a critical period during which interactions with the genital tract microbiota likely contribute to both maternal and fetal outcomes.

A total of 36 women with PPROM were enrolled in the study. However, initial samples for 12 of these women were collected >24 hours following membrane rupture and after the administration of antibiotics, which left 24 T_0 samples for analysis in addressing the question of whether the vaginal microbiome structure at time of PPROM

predicts latency duration. Hierarchical clustering of T_0 profiles based on sequences comprising at least 10% of the microbial profile of at least one woman did not reveal any associations between microbiome profile and gestational age at PPROM, latency duration, or development of chorioamnionitis (Figure 5.2). The results of the study do not provide support for the vaginal microbiome profile at time of membrane rupture as a predictor of latency duration; however the analysis of weekly samples collected during the latency period revealed a remarkably unstable, abnormal microbiome.

BV has been identified as a risk factor for preterm birth and PPROM (Epstein *et al.* 1998; Hillier *et al.* 1995). In the current study, none of the samples collected within 24 hours of membrane rupture (T_0 samples) had Nugent scores consistent with BV, including four that had no cells on the slide. Amniotic fluid leaking after membrane rupture may have dislodged some of the resident microbiota, resulting in reduced numbers of bacteria available for scoring. Lower 16S rRNA copy numbers detected per vaginal swab in T_0 samples in comparison to those detected in swabs from healthy pregnant women supports this explanation. It is also possible that amniotic fluid increased the pH of the vaginal sample collected for Gram staining, which can decrease the efficacy of the stain (Lamanna and Mallette 1954). Samples that did not have sufficient numbers of cells for Nugent scoring did provide adequate genomic DNA template for *cpn*60 universal target profiling, illustrating the relatively low limit of detection of PCR compared to microscopy. A lack of consistency of T_0 microbiota with BV was also apparent in the sequence-based profiles, with few of the microbiome profiles of T_0 samples dominated by *G. vaginalis* (5/24) or *Atopobium vaginae* (0/24), which are strongly associated with a Nugent score diagnosis of BV (Figure 5.2) (Menard *et al.* 2008; Romero *et al.* 2014a).

BVAB (BV associated bacteria) are another group of organisms, recently described as being associated with BV (Fredricks *et al.* 2005; Oakley *et al.* 2008). Three OTUs were detected with BVAB3 and BVAB2 as nearest neighbour, but no samples in this study were dominated by these sequences, with no sample containing more than 15.7% BVAB.

Only half of the T_0 samples had profiles dominated by *Lactobacillus* (Figure 2). Among these, samples dominated by *L. crispatus* (n=4) were outnumbered by those dominated by *L. iners,* either alone or in combination with *L. jensenii.* In fact, *L. iners* was detected in all T_0 samples, although it comprised >50% of the microbiome in only four cases. *L. iners* is associated both with normal and BV microbiota, and is also reported as a dominant organism when the vaginal microbiota is in a transition state, either from normal to BV or vice versa (Jakobsson and Forsum 2007; Santiago *et al.* 2012; Verstraelen *et al.* 2009), and *L. iners* dominated microbiomes have been associated with spontaneous preterm birth (Petricevic *et al.* 2014). The hemolytic and mucinolytic potential of *L. iners* could be involved in its survival during disturbed vaginal conditions, and also may be suggestive of its pathogenic potential (Macklaim *et al.* 2011). In contrast, longitudinal stability of the vaginal microbiota in pregnant women with term deliveries has been associated with *L. crispatus* dominated profiles (Romero *et al.* 2014a). While it is possible that more of the women in the study had *Lactobacillus* dominated microbiota immediately prior to PPROM, these bacteria would have had to have been selectively reduced to extremely low or even undetectable levels in the microbial community due to amniotic fluid leakage or some other factor within 24 hours of membrane rupture.

Prevotella spp. were detected in all samples from all women in this study, and sequences from this genus dominated some profiles either at T_0 or at subsequent time points during latency (Figure 5.2 and Figure 5.6). In contrast, in a *cpn*60-based study of 91 vaginal samples from clinically healthy, non-pregnant, Canadian women, *Prevotella* were detected in only 57% of samples, with the most frequently detected species (*P. timonensis*) present in only 45% of samples (Chaban *et al.* 2014). *Prevotella* have been detected in 62-96% percent of samples in two recent studies of the vaginal microbiota of non-pregnant, reproductive aged women with a range of Nugent scores, but the particular species within this genus could not be resolved using 16S rRNA sequence data (Ravel *et al.* 2011; Shipitsyna *et al.* 2013). *cpn*60 universal target sequences have been demonstrated to provide superior resolution of Gram negative anaerobic rods, including *Bacteroides* and *Prevotella* (Sakamoto and Ohkuma 2010). In the current study, we identified 14 known species and 15 additional, possibly novel *Prevotella* spp. (Figure 5.4). Given the suggested role of *Prevotella* spp. in synergistic relationships with BV associated organisms (Pybus and Onderdonk 1997, 1998) and their production of lipopolysaccharide in the vaginal environment (Aroutcheva *et al.* 2008), the prevalence and diversity of *Prevotella* spp. in the PPROM associated vaginal microbiota clearly warrants further investigation.

OTU sequences corresponding to *Megasphaera* type 1 were also detected in all samples examined. *Megasphaera* spp. were first reported to be part of the vaginal microbiome in 2004 (Zhou *et al.* 2004), and *Megasphaera* type 1 has since been strongly associated with BV (Datcu *et al.* 2013; Zozaya-Hinchliffe *et al.* 2008). Two OTU sequences clustered with *Megasphaera micronuciformis* F0359 (99-100% sequence

identity), an obligate anaerobe isolated from the human oral cavity as a part of the Human Microbiome Project (Figure 5.3). Identification of bacterial species associated with the oral microbiome in the amniotic fluid of women who experience preterm birth has been observed previously, although the significance of this relationship and whether it is indicative of transfer of organisms is not known (Gauthier *et al.* 2011; Hill 1998).

Genital *Mycoplasma* and *Ureaplasma* have been strongly associated with adverse maternal and fetal outcomes (Goldenberg *et al.* 1998; Taylor-Robinson and Lamont 2011; Waites *et al.* 2005), and cervical colonization of genital mycoplasmas is associated with preterm birth and PPROM (Kacerovsky *et al.* 2009; Wasiela *et al.* 2003). These organisms have also been isolated from amniotic fluid and chorioamnion of women with PPROM and preterm labor (Kim *et al.* 2009; Wang *et al.* 2013). Since *cpn*60 genes are known to be absent in some Mollicutes (Clark and Tillier 2010), genus and species specific PCR was used to detect these organisms. Eighty percent of women in the study were found to be PCR positive for *Mycoplasma* and/or *Ureaplasma*, which is at the higher end of the range reported in pregnant women (10-84%) (Bayraktar *et al.* 2010; Grattard *et al.* 1995; Larsen and Hwang 2010; Paul *et al.* 1998). No *M. genitalium* was detected in any of the women in the study, either by pyrosequencing of Mollicutes 16S rRNA amplicons, or by species-specific PCR. *M. genitalium* is strongly associated with urethritis in men and cervicitis and infertility issues in women (Larsen and Hwang 2010). However, its role seems to be more of a cervicitis pathogen (Gaydos *et al.* 2009; Pepin *et al.* 2005) contributing to infertility similar to *Chlamydia trachomatis* rather than as part of an endogenous microbiota contributing to adverse pregnancy outcomes. Hence in some studies, presumably in populations with low rates of cervicitis, they either did not

detect the species in amniotic fluid or were not able to demonstrate any associations with adverse pregnancy outcomes (Cazanave *et al.* 2012; Larsen and Hwang 2010). In the current study, the presence of vaginal Mollicutes was associated with significantly lower APGAR scores, lower GA at delivery, and correspondingly lower birth weight, supporting previous suggestions of a potential role for these organisms in adverse pregnancy outcomes. However, small sample size precluded a statistical analysis of the association of individual species of Mollicutes with these outcomes.

Antibiotic administration in women with PPROM at less than 32 weeks has been shown to decrease neonatal morbidity and prolong pregnancy (Yudin *et al.* 2009), and all women in this study were treated with antimicrobials according to current clinical guidelines. All women were also administered corticosteroids: betamethasone and/or dexamethasone, for the purposes of promoting fetal lung maturation. Antibiotic and steroid administration may have contributed to the microbiome profiles observed meeting the definition of mixed abnormal or intermediate vaginal microbiota or even aerobic vaginitis (*E. coli* dominated) (Donati *et al.* 2010; Donders 2007; Donders *et al.* 2002), rather than *Lactobacillus* dominated or BV. Another potential contributor to the observed abnormal microbiota is the leakage of amniotic fluid (pH 7 to 7.5) (Davidson 1991) through the normally acidic vaginal environment. These disruptive factors likely contributed to the observation that women whose vaginal microbiota was dominated by *Lactobacillus* immediately following PPROM (women 29, 34, 42, 48, 51 and 60, Figure 5.6), rapidly transitioned to an abnormal microbial profile.

Perhaps the most striking observation in this study was the highly unstable nature of the vaginal microbiota over the latency period (Figure 5.6, 5.7). The results of the

graphical analysis of the microbial profiles and the quantitative assessment of ecological distance between samples within and between women are in stark contrast to the generally stable microbial community observed in longitudinal studies of healthy, non-pregnant (Chaban *et al.* 2014; Gajer *et al.* 2012; Hickey *et al.* 2012) and pregnant women (Romero *et al.* 2014a). The current study did not assess the stability of the vaginal microbiota in these women prior to PPROM, and so it cannot be determined if as a group they had a more or less stable microbiota prior to PPROM. It has been suggested that stability in terms of maintenance of a core set of metabolic capabilities of the human microbiome is correlated to health, even when the community structure varies in its taxonomic composition (Consortium 2012). The instability reported in this study is in terms of phylogenetic or taxonomic composition, not functional capacity. However, if stability is an important factor for health, then the highly unstable microbiome observed following PPROM needs to be studied further using metagenomic and metabolomics approaches, to gain a more complete understanding of the relationship of microbiome dynamics and maternal and fetal outcomes, especially long term outcomes following membrane rupture.

In 12 cases, the neonatal skin microbiome was characterized and compared to the vaginal sample closest to delivery (Figure 5.8). Detection of vaginal microbiota on infant skin is hardly surprising in cases of vaginal delivery, but similar organisms were also detected in the skin microbiomes of three infants delivered by C-section. This could be due to ascension of bacteria from the lower genital tract either before or after membrane rupture. Vaginal bacteria such as *Gardnerella*, *Megasphaera*, *Prevotella*, *Porpyhromonas*, and *Ureaplasma* have been recovered from the meconium of infants

delivered either by C-section or vaginally, suggesting possible exposure of the fetus to the maternal microbiota during gestation as well as during delivery (Datcu *et al.* 2013).

Taken together, the results of this study provide an unprecedented characterization of the vaginal microbiota following PPROM, and support previous suggestions of a role for Mollicutes in adverse pregnancy outcomes associated with PPROM. These results also clearly demonstrate the value of application of culture-independent methods for microbial profiling, especially where the nature of the available samples precludes application of techniques such as microscopy and culture, or severely limits their utility. The highly unstable and abnormal vaginal microbiota of the women in this study demonstrates the need for more intense study of the relationship of genital tract microbiota with PPROM, including functional analysis of these microbial communities. Future work should involve larger studies including sampling before and after membrane rupture, and also after delivery, to determine if the composition of the microbiota is abnormal or unstable prior to membrane rupture, and whether the community achieves a healthy state following cessation of antibiotic therapy.

Acknowledgements

This work was funded by a March of Dimes Prematurity Research Initiative grant awarded to DMM [MONEY 21-FY08-537]. TPJ was supported by a University of Saskatchewan Graduate Scholarship. The authors are grateful to Bonnie Chaban for helpful discussion and advice in data analysis. The authors wish to thank the clinicians and laboratory staff at BC Women's Hospital and Health Centre and their patients for their important contributions to this study.

CHAPTER 6 - General Discussion and Conclusions

6.1 Summary and limitations of these works

Gardnerella vaginalis can be resolved into four subgroups that likely represent different species.

Using *cpn*60 UT sequence, *G. vaginalis* could be resolved into four subgroups. Upon inspection of the whole genome sequences of isolates of *G. vaginalis* of four subgroups using inter-genome similarity comparing algorithm ANIm, it was found that the four groups meet the suggested criterion for a molecular definition of bacterial species (Richter and Rossello-Mora 2009). The whole genome based analysis reported by Ahmed *et al.* (2012) also describes four clades within the genus *Gardnerella* and results presented in this thesis have shown that the four clades described by Ahmed *et al* are the same as the four *cpn*60 subgroups that I have described. Proteomic profiles of isolates of each subgroup formed unambiguous clusters supporting the *cpn*60 UT based and Ahmed *et al.* based categorization of the genus.

Determination of the Piot biotypes (1984) of *G. vaginalis* isolates was not useful in defining subgroups as there was no congruence between *cpn*60 subgroups and Piot biotypes. However, *ß*-galactosidase activity, one of the phenotyping assays included in Piot biotyping, was found to be only exhibited by subgroup C *G. vaginalis* isolates. Similarly, antimicrobial susceptibility testing of available *G. vaginalis* isolates was not useful in discriminating subgroups.

One of the limitations of this part of research is that with more *cpn*60 sequences being added for the phylogenetic analysis presented in Chapter 3 than were included in

our original study (Chapter 2), an outlier group has been identified. Our current culture collection of *G. vaginalis* does not include any isolates belonging to this outlier group and so I was unable to investigate properties of these isolates. One of the other drawbacks is that there is only one isolate in our culture collection from subgroup D, which limits my ability to draw strong conclusions about this subgroup. A previous metagenomic study (Chaban *et al.* 2014) of vaginal microbiome of healthy reproductive aged women indicates that subgroup D is rare. Subgroup D *G. vaginalis* is also observed to be rare in healthy reproductive aged pregnant and non-pregnant Canadian women (n=500) profiled as part of the ongoing work of our research team (Vogue Research Team, Unpublished). Collecting more isolates of this subgroup and investigating phenotypic and genotypic properties will help provide evidence to support likelihood of subgroups being different species.

Clinical significance of subgroup B *Gardnerella vaginalis* is demonstrated by the presence of sialidase enzymatic activity and its differential distribution in Kenyan women

Subgroup B *G. vaginalis* stands out from the rest of the subgroups in terms of a few clinically significant factors. Sialidase activity, which is associated strongly with BV, is one characteristic that can be used to discriminate subgroup B from the rest. Another observation of clinical significance comes from the observation of differential distribution of subgroup B isolates in a cohort of Kenyan women. Subgroup B *G. vaginalis* was significantly more abundant in women categorized as having BV than not. Also, the fascinating observation of anaerobic isolates of subgroup B also makes them interesting and important candidates for consideration while developing better diagnosis of BV.

While describing the clinical significance of subgroup B, it is important to be cautious as the distribution of *G. vaginalis* subgroups in many more cohorts needs to be understood before concluding anything about the clinical significance of this subgroup. Similarly, sialidase is so far the only virulence factor that distinguishes subgroup B from others. There are many other important virulence factors such as prolidase and vaginolysin that need to be further characterized in isolates within the four subgroups. This in turn will help understand any other factors that might distinguish subgroups as well as improving understanding of the clinical significance associated with these subgroups. The potential clinical significance of anaerobic isolates of subgroup B *G. vaginalis* needs to be explored further given the suggested role of *G. vaginalis* as a "stage setter" for BV.

One of the major limitations of this work is that genome content differences between anaerobic and facultative strains of *G. vaginalis* have not been characterized beyond predicting functions based on genome annotation. These genomic fragments require additional characterization and validation in terms of function. In addition to that, both isolates that exhibited an anaerobic phenotype (N153 and N101) were isolated from the same woman. Though it is reported in the literature that anaerobic isolates of *Gardnerella* are rare, further efforts are required to isolate anaerobic strains of *Gardnerella* from different women, possibility even from different cohorts. Anaerobic isolates of *G. vaginalis* were only detected in subgroup B, but more isolates of other subgroups need to be characterized and investigated for the presence of anaerobic isolates to determine if anaerobic lifestyle is a subgroup B specific characteristic.

Overall microbial profiles at membrane rupture in women with PPROM did not predict clinical outcomes

The prospective study of the vaginal microbiota following preterm premature rupture of membranes resulted in an unprecedented characterization of the vaginal microbial community structure of these women immediately after membrane rupture and also during their latency period. Skin microbiomes of neonates born to these women were also characterized and described in terms of species present and their adundance. Although the vaginal microbiome has been studied numerous times by researchers around the world for the past many decades, the vaginal microbial community after membrane rupture has to our knowledge never been described. Characterizing the vaginal microbiome immediately after membrane rupture, and through latency until delivery is a great addition to the wealth of information that we have specifically about the condition of PPROM, and more information regarding the same might help clinicians better manage this potentially devastating condition.

It was found that microbiome profiles immediately after membrane rupture were generally dominated by one or two species including *Lactobacillus crispatus, L. iners, Prevotella timonensis, Prevotella bivia, G. vaginalis*, and *Escherichia coli.* The vaginal microbiota was found to be highly unstable over the latency period, with dramatic shifts in composition between weekly samples. There was an overall decrease in *Lactobacillus* abundance as weeks advanced in the latency period. Mollicutes (*Mycoplasma* and/or *Ureaplasma*) were detected by conventional PCR in 81% of women. Adverse pregnancy outcome were correlated with detection of Mollicutes as these women with Mollicutes

had significantly lower GA at delivery and correspondingly lower birth weight infants than Mollicutes PCR negative women.

This prospective study has helped generate many testable hypotheses that should be followed up on future studies. One such hypothesis is that *Megasphaera* and *Prevotella* comprise a core microbiome in women with PPROM. Higher abundance and prevalence of *Lactobacillus iners* in these women and its association with microbiome instability and transition needs to be carefully questioned in future studies.

One of the major limitations of this study was the sample number. Only 36 women were included in the microbiome analyses, a number that does not power us statistically to make strong correlations and hence hampers making concrete conclusions. It remains to be determined if the lack of prediction of latency duration based on microbiome profile at time of membrane rupture was due to the study being underpowered. This cohort presents many challenges in terms of patient enrolment and sampling, and we experienced many sampling problems, both in terms of having few samples being collected (low study enrolment rates) and insufficient quantities of amplifiable DNA in each sample (which is likely due to having collected vaginal samples after the flush out of microbial community by the amniotic fluid following membrane rupture). At the end of enrolment, we had 62 women enrolled for the study. However, due to issues with previously used sampling protocols based on the DNAzol Direct method, samples from 22 women were rendered useless as they had insufficient quantities of amplifiable DNA. Adoption of new sample collection and processing protocol based on using a magnetic bead-based kit resulted in saving the rest of the samples.

6.2 Discussion of future prospects

Species level resolution required to make clinical associations

A species in the microbial world is currently defined by a polyphasic approach based on genotypic and phenotypic properties using techniques from microbial systematics, ecology, genomics and evolutionary biology. Species resolution is critical in microbiome studies where microbiome structure and patterns associated with clinical outcomes could lead to important diagnostic procedures. Average nucleotide identity based on whole genome sequences is replacing the gold standard method for defining species, which is DNA-DNA hybridization. In metagenomic studies, microbial diversity is usually described using operational taxonomic units (OTUs) based on 16S rRNA gene sequences. However, 16S rRNA provides only limited resolution and helps in discriminating microbes in many cases only to genus level. The *cpn*60 universal target is demonstrated to provide high-resolution microbiome profiles (Chaban *et al.* 2014; Hill *et al.* 2005a; Schellenberg *et al.* 2011a; Schellenberg *et al.* 2009b) and has been proven to be an effective tool in predicting genome similarity and for species determination (Verbeke *et al.* 2011).

In Chapters 2 and 3, I have demonstrated the ability of *cpn*60 to resolve the heterogeneity within the genus *Gardnerella*. There are other taxa that may benefit from closer examination and potential sub-speciation using these methods. Examples of such taxa include *Megasphaera* and the BVABs since their positive association with BV and high specificity for diagnosing BV make them very important taxa that require urgent

investigation. Currently, *Megasphaera* are rarely reported at higher resolution than genus, and the lack of cultured isolates of BVAB1 and BVAB2 has meant that they are reported solely on the basis of 16S rRNA sequence detection. In some instances, resolution greater than species is a requirement. For instance, *Lactobacillus iners* is an important taxon of vaginal microbiota, noted for its presence when the microbiome is in transition state. *L. iners* is also routinely detected in the healthy microbiome and is often a dominant member of healthy vaginal microbiota. In the PPROM cohort, two OTUs corresponding to *L. iners* were detected. Such strain level resolution needs to be further inspected in other cohorts and studied to understand any clinical significance that might be present for the different *L. iners* strains.

Unexplored health states needs to be studied

The vaginal microbiota of adolescence, reproductive age, and post-menopause has been well described using cultivation based methods as well as molecular methods. The vaginal microbiota during pregnancy has also been described as being less diverse and less rich in comparison to reproductive aged non-pregnant women. However, we still do not understand the role of the vaginal microbiota in the infectious etiology of preterm birth, PPROM and other adverse pregnancy outcomes. Our study was designed specifically to address the question whether microbiome at PPROM can predict maternal and fetal outcomes. Our study was not designed to conclude whether the microbiome that we see after PPROM is a cause or an effect of PPROM. It is also difficult to state the reason behind the unstable nature of microbiome during latency, beyond the obvious contribution of antibiotic use. However, these questions are critical in improving the understanding of clinicians to manage the condition of PPROM.

To fill these gaps in knowledge, it is critical to understand the composition and the stability of the postpartum microbiome. Delivery induces abrupt and dramatic changes in levels of hormones such as progesterone, estrogen, prolactin, cortisol, oxytocin, thyroid, and vasopressin. In addition to that, vaginal pH changes abruptly due to the flow of amniotic fluid when membrane rupture occurs prior to delivery. Literature describing how hormonal factors and other changes during delivery affects the vaginal microbiota and its stability postpartum is limited. The role of host genetics in the vaginal ecosystem is still unexplored. We not only need to understand the effects of varying health parameters on the vaginal microbiota, but also need to be able to address how the metabolic and immunologic components of the vaginal ecosystem change with respect to different health status.

Better diagnostic methods need to replace traditional diagnostic methods

The gold standard method of diagnosing abnormal vaginal microbiota is the Nugent score, which is a weighted scoring system based on abundance of bacterial morphotypes defined by Gram stain score. There are many limitations associated with this traditional diagnostic method. As already stated, Nugent score might vary based on the technical staff scoring the slide. The method is semi-quantitative and low resolution. Bacterial morphotypes, recognized as a part of Nugent scoring, has been largely identified from data inferred from cultivation-based studies. A recent study (Srinivasan *et al.* 2013) that compared bacterial morphotypes recognized using Nugent scoring method to bacteria detected using high-throughput sequencing methods indicated that Gram-negative rods designated *Mobiluncus* morphotypes on Gram stain are in fact BVAB1.

Srinivasan *et al.* (2013) demonstrated the presence of BVAB1 and also its abundance through microscopic evidence using fluorescence *in situ* hybridization (FISH). Novel species such as BVAB1, BVAB2, *Leptotrichia* spp., *Atopobium* spp., and *Megasphaera* spp., not included in the bacterial morphotypes that define Nugent score, are now being repeatedly reported as being highly specific for diagnosing BV. As we have seen, Nugent scoring can be rendered useless in some cases of women with PPROM as slides do not contain cells. In such cases, we require an alternate diagnostic method to describe the kind of microbiota present in women to determine the kind of antimicrobial treatment that needs to be administered.

Virulence factors such as sialidase and prolidase are known to be predictors of BV. In addition to this, metabolic components in the vagina such as trimethyl amine are also described to strongly predict BV. Hence, developing a composite indicator variable combining the microbiome, virulence factors, and metabolic components might serve as a biomarker and could serve as better diagnostic method.

To conclude, research needs to focus on generating high resolution profiles in combination with other aspects affecting health, considering microbial communities as part of human ecosystem. The concept of describing individual species as "pathogens" should be constrained, and these organisms should be considered within the context of the community as a whole. One needs to remember that in an ecosystem with a diverse microbial community, there will be variance associated with the metabolic and virulence potential based on the type of microbes in the community and parameters associated with the host. If we are able to define health using the above described principle and are able

to discriminate healthy from diseased states, we will be successful in generating better diagnostic methods that would help clinicians manage diseased states better.

REFERENCES

1988. International committee on systematic bacteriology announcement of the report of head hoc committee on reconciliation of approaches to bacterial systematics. Blackwell Publishing Ltd. 00218847.

2002. Sexually transmitted diseases treatment guidelines 2002. Centers for Disease Control and Prevention. Guideline Practice Guideline. 1057-5987 (Print) 1057-5987 (Linking).

2007. ACOG Practice Bulletin No. 80: premature rupture of membranes. Clinical management guidelines for obstetrician-gynecologists. Practice Guideline. 0029-7844 (Print) 0029-7844 (Linking).

Aalto, M., Ronnemaa, T., and Kulonen, E. 1974. Effect of neuraminidase on the synthesis of collagen and other proteins. Biochimica et Biophysica Acta **342**(2): 247-253.

Ahmed, A., Earl, J., Retchless, A., Hillier, S.L., Rabe, L.K., Cherpes, T.L., Powell, E., Janto, B., Eutsey, R., Hiller, N.L., Boissy, R., Dahlgren, M.E., Hall, B.G., Costerton, J.W., Post, J.C., Hu, F.Z., and Ehrlich, G.D. 2012. Comparative genomic analyses of 17 clinical isolates of *Gardnerella vaginalis* provide evidence of multiple genetically isolated clades consistent with subspeciation into genovars. Journal of Bacteriology **194**(15): 3922-3937.

Amsel, R., Totten, P.A., Spiegel, C.A., Chen, K.C., Eschenbach, D., and Holmes, K.K. 1983. Nonspecific vaginitis. Diagnostic criteria and microbial and epidemiologic associations. American Journal of Medicine **74**(1): 14-22.

Aroutcheva, A., Ling, Z., and Faro, S. 2008. *Prevotella bivia* as a source of lipopolysaccharide in the vagina. Anaerobe **14**(5): 256-260.

Aroutcheva, A.A., Simoes, J.A., Behbakht, K., and Faro, S. 2001. *Gardnerella vaginalis* isolated from patients with bacterial vaginosis and from patients with healthy vaginal ecosystems. Clinical Infectious Diseases **33**(7): 1022-1027.

Bailey, R.K., Voss, J.L., and Smith, R.F. 1979. Factors affecting isolation and identification of *Haemophilus vaginalis* (*Corynebacterium vaginale*). Journal of Clinical Microbiology **9**(1): 65-71.

Balashov, S.V., Mordechai, E., Adelson, M.E., and Gygax, S.E. 2014. Identification, quantification and subtyping of *Gardnerella vaginalis* in noncultured clinical vaginal samples by quantitative PCR. Journal of Medical Microbiology **63**(Pt 2): 162-175.

Bartlett, J.G., Onderdonk, A.B., Drude, E., Goldstein, C., Anderka, M., Alpert, S., and McCormack, W.M. 1977. Quantitative bacteriology of the vaginal flora. Journal of Infectious Diseases **136**(2): 271-277.

Bayraktar, M.R., Ozerol, I.H., Gucluer, N., and Celik, O. 2010. Prevalence and antibiotic susceptibility of *Mycoplasma hominis* and *Ureaplasma urealyticum* in pregnant women. International Journal of Infectious Diseases **14**(2): e90-95.

Benito, R., Vazquez, J.A., Berron, S., Fenoll, A., and Saez-Neito, J.A. 1986. A modified scheme for biotyping *Gardnerella vaginalis*. Journal of Medical Microbiology **21**(4): 357-359.

Blaiotta, G., Fusco, V., Ercolini, D., Aponte, M., Pepe, O., and Villani, F. 2008. *Lactobacillus* strain diversity based on partial hsp60 gene sequences and design of PCR-restriction fragment length polymorphism assays for species identification and differentiation. Applied and Environmental Microbiology **74**(1): 208-215.

Boskey, E.R., Cone, R.A., Whaley, K.J., and Moench, T.R. 2001. Origins of vaginal acidity: high D/L lactate ratio is consistent with bacteria being the primary source. Human Reproduction **16**(9): 1809-1813.

Briselden, A.M., Moncla, B.J., Stevens, C.E., and Hillier, S.L. 1992. Sialidases (neuraminidases) in bacterial vaginosis and bacterial vaginosis-associated microflora. Journal of Clinical Microbiology **30**(3): 663-666.

Brotman, R.M., Shardell, M.D., Gajer, P., Fadrosh, D., Chang, K., Silver, M.I., Viscidi, R.P., Burke, A.E., Ravel, J., and Gravitt, P.E. 2014. Association between the vaginal microbiota, menopause status, and signs of vulvovaginal atrophy. Menopause **21**(5): 450-458.

Brousseau, R., Hill, J.E., Prefontaine, G., Goh, S.H., Harel, J., and Hemmingsen, S.M. 2001. *Streptococcus suis* serotypes characterized by analysis of chaperonin 60 gene sequences. Applied and Environmental Microbiology **67**(10): 4828-4833.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., and Knight, R. 2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods **7**(5): 335-336.

Catlin, B.W. 1992. *Gardnerella vaginalis*: characteristics, clinical considerations, and controversies. Clinical Microbiology Reviews **5**(3): 213-237.

Cauci, S., and Culhane, J.F. 2011. High sialidase levels increase preterm birth risk among women who are bacterial vaginosis-positive in early gestation. American Journal of Obstetrics and Gynecology **204**(2): 142 e141-149.

Cauci, S., Guaschino, S., Driussi, S., De Santo, D., Lanzafame, P., and Quadrifoglio, F. 2002. Correlation of local interleukin-8 with immunoglobulin A against *Gardnerella vaginalis* hemolysin and with prolidase and sialidase levels in women with bacterial vaginosis. Journal of Infectious Diseases **185**(11): 1614-1620.

Cauci, S., McGregor, J., Thorsen, P., Grove, J., and Guaschino, S. 2005. Combination of vaginal pH with vaginal sialidase and prolidase activities for prediction of low birth weight and preterm birth. American Journal of Obstetrics and Gynecology **192**(2): 489-496.

Cauci, S., Monte, R., Ropele, M., Missero, C., Not, T., Quadrifoglio, F., and Menestrina, G. 1993. Pore-forming and haemolytic properties of the *Gardnerella vaginalis* cytolysin. Molecular Microbiology **9**(6): 1143-1155.

Cazanave, C., Manhart, L.E., and Bebear, C. 2012. *Mycoplasma genitalium*, an emerging sexually transmitted pathogen. Médecine et Maladies Infectieuses **42**(9): 381-392.

Chaban, B., Albert, A., Links, M.G., Gardy, J., Tang, P., and Hill, J.E. 2013. Characterization of the upper respiratory tract microbiomes of patients with pandemic H1N1 influenza. PLoS ONE **8**(7): e69559.

Chaban, B., Links, M.G., and Hill, J.E. 2012. A molecular enrichment strategy based on cpn60 for detection of epsilon-proteobacteria in the dog fecal microbiome. Microbial Ecology **63**(2): 348-357.

Chaban, B., Links, M.G., Jayaprakash, T.P., Wagner, E.C., Bourque, D.K., Lohn, Z., Albert, A.Y., van Schalkwyk, J., Reid, G., and Hemmingsen, S.M. 2014. Characterization of the vaginal microbiota of healthy Canadian women through the menstrual cycle. Microbiome **2**(1): 23.

Chaban, B., Musil, K.M., Himsworth, C.G., and Hill, J.E. 2009. Development of cpn60-based real-time quantitative PCR assays for the detection of 14 *Campylobacter* species and application to screening canine fecal samples. Applied and Environmental Microbiology **75**(10): 3055-3061.

Cho, J.C., and Tiedje, J.M. 2001. Bacterial species determination from DNA-DNA hybridization by using genome fragments and DNA microarrays. Applied and Environmental Microbiology **67**(8): 3677-3682.

Christopher, P.H., Mario, M., Joy, E.L., Jennifer, H.R., and Lale, S. 2009. White Paper on Preterm Birth: The Global and Regional Toll. The March of Dimes Foundation.

Clark, G.W., and Tillier, E.R. 2010. Loss and gain of GroEL in the Mollicutes. Biochemistry and Cell Biology **88**(2): 185-194.

Cohan, F.M. 2002. What are bacterial species? Annual Review of Microbiology **56**: 457-487.

Cole, A.M. 2006. Innate host defense of human vaginal and cervical mucosae. Current Topics in Microbiology and Immunology **306**: 199-230.

Consortium, T.H.M.P. 2012. Structure, function and diversity of the healthy human microbiome. Nature **486**(7402): 207-214.

Cook, R.L., Hutchison, S.L., Ostergaard, L., Braithwaite, R.S., and Ness, R.B. 2005. Systematic review: noninvasive testing for *Chlamydia trachomatis* and *Neisseria gonorrhoeae*. Annals of Internal Medicine **142**(11): 914-925.

Cruden, D.L., and Galask, R. 1988. Reduction of trimethylamine oxide to trimethylamine by *Mobiluncus* strains isolated from patients with bacterial vaginosis. Microbial Ecology in Health and Disease **1**(2): 95-100.

Datcu, R., Gesink, D., Mulvad, G., Montgomery-Andersen, R., Rink, E., Koch, A., Ahrens, P., and Jensen, J.S. 2013. Vaginal microbiome in women from Greenland assessed by microscopy and quantitative PCR. BMC Infectious Diseases **13**: 480.

Davidson, K.M. 1991. Detection of premature rupture of the membranes. Clinical Obstetrics and Gynecology **34**(4): 715-722.

Demba, E., Morison, L., van der Loeff, M.S., Awasana, A.A., Gooding, E., Bailey, R., Mayaud, P., and West, B. 2005a. Bacterial vaginosis, vaginal flora patterns and vaginal hygiene practices in patients presenting with vaginal discharge syndrome in The Gambia, West Africa. BMC Infectious Diseases **5**: 12.

Demba, E., Morison, L., van der Loeff, M.S., Awasana, A.A., Gooding, E., Bailey, R., Mayaud, P., and West, B. 2005b. Bacterial vaginosis, vaginal flora patterns and vaginal hygiene practices in patients presenting with vaginal discharge syndrome in the Gambia, West Africa. BMC Infectious Diseases **5**(1): 12.

Desai, A.R., Links, M.G., Collins, S.A., Mansfield, G.S., Drew, M., Van Kessel, A.G., and Hill, J.E. 2012. Effects of plant-based diets on the distal gut microbiome of rainbow trout (*Oncorhynchus mykiss*). Aquaculture **350-353**: 134-142.

Desai, A.R., Musil, K.M., Carr, A.P., and Hill, J.E. 2009. Characterization and quantification of feline fecal microbiota using cpn60 sequence-based methods and investigation of animal-to-animal variation in microbial population structure. Veterinary Microbiology **137**(1-2): 120-128.

DiGiulio, D.B., Romero, R., Kusanovic, J.P., Gomez, R., Kim, C.J., Seok, K.S., Gotsch, F., Mazaki-Tovi, S., Vaisbuch, E., Sanders, K., Bik, E.M., Chaiworapongsa, T., Oyarzun, E., and Relman, D.A. 2010. Prevalence and diversity of microbes in the amniotic fluid, the fetal inflammatory response, and pregnancy outcome in women with preterm pre-labor rupture of membranes. American Journal of Reproductive Immunology **64**(1): 38-57.

Donati, L., Di Vico, A., Nucci, M., Quagliozzi, L., Spagnuolo, T., Labianca, A., Bracaglia, M., Ianniello, F., Caruso, A., and Paradisi, G. 2010. Vaginal microbial flora and outcome of pregnancy. Archives of Gynecology and Obstetrics **281**(4): 589-600.

Donders, G.G. 2007. Definition and classification of abnormal vaginal flora. Bailliere's Best Practice and Research. Clinical Obstetrics and Gynaecology **21**(3): 355-373.

Donders, G.G., Vereecken, A., Bosmans, E., Dekeersmaecker, A., Salembier, G., and Spitz, B. 2002. Definition of a type of abnormal vaginal flora that is distinct from bacterial vaginosis: aerobic vaginitis. BJOG **109**(1): 34-43.

Dover, S.E., Aroutcheva, A.A., Faro, S., and Chikindas, M.L. 2008. Natural antimicrobials and their role in vaginal health: a short review. International Journal of Probiotics and Prebiotics **3**(4): 219-230.

Dumonceaux, T.J., Hill, J.E., Hemmingsen, S.M., and Van Kessel, A.G. 2006a. Characterization of intestinal microbiota and response to dietary virginiamycin supplementation in the broiler chicken. Applied and Environmental Microbiology **72**(4): 2815-2823.

Dumonceaux, T.J., Hill, J.E., Pelletier, C., Paice, M.G., Van Kessel, A.G., and Hemmingsen, S.M. 2006b. Molecular characterization of microbial communities in Canadian pulp and paper activated sludge and quantification of a novel *Thiothrix eikelboomii*-like bulking filament. Canadian Journal of Microbiology **52**(5): 494-500.

Dumonceaux, T.J., Schellenberg, J., Goleski, V., Hill, J.E., Jaoko, W., Kimani, J., Money, D., Ball, T.B., Plummer, F.A., and Severini, A. 2009. Multiplex detection of bacteria associated with normal microbiota and with bacterial vaginosis in vaginal swabs by use of oligonucleotide-coupled fluorescent microspheres. Journal of Clinical Microbiology **47**(12): 4067-4077.

Edwards, D.I. 1993. Nitroimidazole drugs--action and resistance mechanisms. II. Mechanisms of resistance. Journal of Antimicrobial Chemotherapy **31**(2): 201-210.

Egan, M.E., and Lipsky, M.S. 2000. Diagnosis of vaginitis. American Family Physician **62**(5): 1095-1104.

Ehrmann, M., Ehrle, R., Hofmann, E., Boos, W., and Schlösser, A. 1998. The ABC maltose transporter. Molecular Microbiology **29**(3): 685-694.

Epstein, F.H., Parry, S., and Strauss, J.F. 1998. Premature rupture of the fetal membranes. New England Journal of Medicine **338**(10): 663-670.

Eschenbach, D.A., Thwin, S.S., Patton, D.L., Hooton, T.M., Stapleton, A.E., Agnew, K., Winter, C., Meier, A., and Stamm, W.E. 2000. Influence of the normal menstrual cycle on vaginal tissue, discharge, and microflora. Clinical Infectious Diseases **30**(6): 901-907.

Farage, M., and Maibach, H. 2006. Lifetime changes in the vulva and vagina. Archives of Gynecology and Obstetrics **273**(4): 195-202.

Farage, M.A., and Maibach, H.I. 2011. Morphology and physiological changes of genital skin and mucosa. Current Problems in Dermatology **40**: 9-19.

Felsenstein, J. 1993. PHYLIP (Phylogeny Infererence Package) version 3.5c, distributed by the author, Department of Genetics, University of Washington, Seattle, USA.

Felsenstein, J. 1996. Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. Methods in Enzymology **266**: 418-427.

Fernando, W.M.U., Hill, J.E., Zello, G.A., Tyler, R.T., Dahl, W.J., and Van Kessel, A.G. 2010. Diets supplemented with chickpea or its main oligosaccharide component raffinose modify fecal microbial composition in healthy adults. Beneficial Microbes **1**(2): 197-207.

Ferris, D.G., Litaker, M.S., Woodward, L., Mathis, D., and Hendrich, J. 1995. Treatment of bacterial vaginosis: a comparison of oral metronidazole, metronidazole vaginal gel, and clindamycin vaginal cream. Journal of Family Practice **41**(5): 443-449.

Fethers, K.A., Fairley, C.K., Hocking, J.S., Gurrin, L.C., and Bradshaw, C.S. 2008. Sexual risk factors and bacterial vaginosis: a systematic review and meta-analysis. Clinical Infectious Diseases **47**(11): 1426-1435.

Fierer, N., Lauber, C.L., Zhou, N., McDonald, D., Costello, E.K., and Knight, R. 2010. Forensic identification using skin bacterial communities. Proceedings of the National Academy of Sciences **107**(14): 6477-6481.

Fortner, K.B., Grotegut, C.A., Ransom, C.E., Bentley, R.C., Feng, L., Lan, L., Heine, R.P., Seed, P.C., and Murtha, A.P. 2014. Bacteria localization and chorion thinning among preterm premature rupture of membranes. PLoS ONE **9**(1): e83338.

Foxman, B., Wen, A., Srinivasan, U., Goldberg, D., Marrs, C.F., Owen, J., Wing, D.A., and Misra, D. 2013. *Mycoplasma*, bacterial vaginosis-associated bacteria BVAB3, race, and risk of preterm birth in a high-risk cohort. American Journal of Obstetrics and Gynecology.

Fredricks, D.N., Fiedler, T.L., and Marrazzo, J.M. 2005. Molecular identification of bacteria associated with bacterial vaginosis. New England Journal of Medicine **353**(18): 1899-1911.

Fredricks, D.N., Fiedler, T.L., Thomas, K.K., Oakley, B.B., and Marrazzo, J.M. 2007. Targeted PCR for detection of vaginal bacteria associated with bacterial vaginosis. Journal of Clinical Microbiology **45**(10): 3270-3276.

Frenette, P., Dodds, L., Armson, B.A., and Jangaard, K. 2013. Preterm prelabour rupture of membranes: effect of latency on neonatal and maternal outcomes. Journal of Obstetrics and Gynaecology Canada **35**(8): 710-717.

Gajer, P., Brotman, R.M., Bai, G., Sakamoto, J., Schutte, U.M., Zhong, X., Koenig, S.S., Fu, L., Ma, Z.S., Zhou, X., Abdo, Z., Forney, L.J., and Ravel, J. 2012. Temporal dynamics of the human vaginal microbiota. Science Translational Medicine **4**(132): 132ra152.

Gardner, H.L., and Dukes, C.D. 1955. *Haemophilus vaginalis* vaginitis: a newly defined specific infection previously classified non-specific vaginitis. American Journal of Obstetrics and Gynecology **69**(5): 962-976.

Gauthier, S., Tetu, A., Himaya, E., Morand, M., Chandad, F., Rallu, F., and Bujold, E. 2011. The origin of *Fusobacterium nucleatum* involved in intra-amniotic infection and preterm birth. Journal of Maternal-Fetal and Neonatal Medicine **24**(11): 1329-1332.

Gaydos, C., Maldeis, N.E., Hardick, A., Hardick, J., and Quinn, T.C. 2009. *Mycoplasma genitalium* as a contributor to the multiple etiologies of cervicitis in women attending sexually transmitted disease clinics. Sexually Transmitted Diseases **36**(10): 598-606.

Gelber, S.E., Aguilar, J.L., Lewis, K.L., and Ratner, A.J. 2008. Functional and phylogenetic characterization of Vaginolysin, the human-specific cytolysin from *Gardnerella vaginalis*. Journal of Bacteriology **190**(11): 3896-3903.

Gibbs, R.S. 2007. Asymptomatic bacterial vaginosis: is it time to treat? American Journal of Obstetrics and Gynecology **196**(6): 495-496.

Goh, S.H., Facklam, R.R., Chang, M., Hill, J.E., Tyrrell, G.J., Burns, E.C., Chan, D., He, C., Rahim, T., Shaw, C., and Hemmingsen, S.M. 2000. Identification of *Enterococcus* species and phenotypically similar *Lactococcus* and *Vagococcus* species by reverse checkerboard hybridization to chaperonin 60 gene sequences. Journal of Clinical Microbiology **38**(11): 3953-3959.

Goh, S.H., Potter, S., Wood, J.O., Hemmingsen, S.M., Reynolds, R.P., and Chow, A.W. 1996. HSP60 gene sequences as universal targets for microbial species identification: studies with coagulase-negative staphylococci. Journal of Clinical Microbiology **34**(4): 818-823.

Goldenberg, R.L., and Culhane, J.F. 2003. Infection as a cause of preterm birth. Clinics in Perinatology **30**(4): 677-700.

Goldenberg, R.L., Culhane, J.F., Iams, J.D., and Romero, R. 2008. Epidemiology and causes of preterm birth. Lancet **371**(9606): 75-84.

Goldenberg, R.L., Culhane, J.F., and Johnson, D.C. 2005. Maternal infection and adverse fetal and neonatal outcomes. Clinics in Perinatology **32**(3): 523-559.

Goldenberg, R.L., Iams, J.D., Mercer, B.M., Meis, P.J., Moawad, A.H., Copper, R.L., Das, A., Thom, E., Johnson, F., McNellis, D., Miodovnik, M., Van Dorsten, J.P., Caritis, S.N., Thurnau, G.R., and Bottoms, S.F. 1998. The preterm prediction study: the value of new vs standard risk factors in predicting early and all spontaneous preterm births. NICHD MFMU Network. American Journal of Public Health **88**(2): 233-238.

Goldstein, E.J., Citron, D.M., Merriam, C.V., Warren, Y.A., Tyrrell, K.L., and Fernandez, H.T. 2002. In vitro activities of Garenoxacin (BMS 284756) against 108 clinical isolates of *Gardnerella vaginalis*. Antimicrobial Agents and Chemotherapy **46**(12): 3995-3996.

Graham, S., Howes, C., Dunsmuir, R., and Sandoe, J. 2009. Vertebral osteomyelitis and discitis due to *Gardnerella vaginalis*. Journal of Medical Microbiology **58**(Pt 10): 1382-1384.

Grattard, F., Soleihac, B., De Barbeyrac, B., Bebear, C., Seffert, P., and Pozzetto, B. 1995. Epidemiologic and molecular investigations of genital mycoplasmas from women and neonates at delivery. Pediatric Infectious Disease Journal **14**(10): 853-858.

Gravett, M.G., Nelson, H.P., DeRouen, T., Critchlow, C., Eschenbach, D.A., and Holmes, K.K. 1986. Independent associations of bacterial vaginosis and *Chlamydia trachomatis* infection with adverse pregnancy outcome. The Journal of the American Medical Association **256**(14): 1899-1903.

Greaves, W.L., Chungafung, J., Morris, B., Haile, A., and Townsend, J.L. 1988. Clindamycin versus metronidazole in the treatment of bacterial vaginosis. Obstetrics and Gynecology **72**(5): 799-802.

Green, N.S., Damus, K., Simpson, J.L., Iams, J., Reece, E.A., Hobel, C.J., Merkatz, I.R., Greene, M.F., and Schwarz, R.H. 2005. Research agenda for preterm birth: recommendations from the March of Dimes. American Journal of Obstetrics and Gynecology **193**(3 Pt 1): 626-635.

Greenwood, J.R., and Pickett, M.J. 1979. Salient features of *Haemophilus vaginalis*. Journal of Clinical Microbiology **9**(2): 200-204.

Greenwood, J.R., and Poickett, M.J. 1980. Transfer of *Haemophilus vaginalis* Gardner and Dukes to a new genus, *Gardnerella*: *G. vaginalis* (Gardner and Dukes) comb. nov. International Journal Of Systemic Bacteriology **30**(1): 170-178.

Haggerty, C.L., Totten, P.A., Ferris, M., Martin, D.H., Hoferka, S., Astete, S.G., Ondondo, R., Norori, J., and Ness, R.B. 2009. Clinical characteristics of bacterial vaginosis among women testing positive for fastidious bacteria. Sexually Transmitted Infections **85**(4): 242-248.

Hammerschlag, M.R., Alpert, S., Rosner, I., Thurston, P., Semine, D., McComb, D., and McCormack, W.M. 1978. Microbiology of the vagina in children: normal and potentially pathogenic organisms. Pediatrics **62**(1): 57-62.

Harwich, M.D., Jr., Alves, J.M., Buck, G.A., Strauss, J.F., 3rd, Patterson, J.L., Oki, A.T., Girerd, P.H., and Jefferson, K.K. 2010. Drawing the line between commensal and pathogenic *Gardnerella vaginalis* through genome analysis and virulence studies. BMC Genomics **11**: 375.

Hay, P.E., Lamont, R.F., Taylor-Robinson, D., Morgan, D.J., Ison, C., and Pearson, J. 1994. Abnormal bacterial colonisation of the genital tract and subsequent preterm delivery and late miscarriage. British Medical Journal **308**(6924): 295-298.

Hekstra, D., and Tommassen, J. 1993. Functional exchangeability of the ABC proteins of the periplasmic binding protein-dependent transport systems Ugp and Mal of *Escherichia coli*. Journal of Bacteriology **175**(20): 6546-6552.

Henry, L.G., McKenzie, R.M., Robles, A., and Fletcher, H.M. 2012. Oxidative stress resistance in *Porphyromonas gingivalis*. Future Microbiology **7**(4): 497-512.

Hesselberg, M., and Vreeland, R.H. 1995. Utilization of protein profiles for the characterization of halophilic bacteria. Current Microbiology **31**(3): 158-162.

Hickey, R.J., Zhou, X., Pierson, J.D., Ravel, J., and Forney, L.J. 2012. Understanding vaginal microbiome complexity from an ecological perspective. Translational Research **160**(4): 267-282.

Higgins, C.F. 1992. ABC transporters: from microorganisms to man. Annual Review of Cell Biology **8**: 67-113.

Higgins, C.F. 2001. ABC transporters: physiology, structure and mechanism--an overview. Research in Microbiology **152**(3-4): 205-210.

Hill, G.B. 1998. Preterm birth: associations with genital and possibly oral microflora. Annals of Periodontology **3**(1): 222-232.

Hill, J.E., Goh, S.H., Money, D.M., Doyle, M., Li, A., Crosby, W.L., Links, M., Leung, A., Chan, D., and Hemmingsen, S.M. 2005a. Characterization of vaginal microflora of healthy, nonpregnant women by chaperonin-60 sequence-based methods. American Journal of Obstetrics and Gynecology **193**(3 Pt 1): 682-692.

Hill, J.E., Hemmingsen, S.M., Goldade, B.G., Dumonceaux, T.J., Klassen, J., Zijlstra, R.T., Goh, S.H., and Van Kessel, A.G. 2005b. Comparison of ileum microflora of pigs fed corn-, wheat-, or barley-based diets by chaperonin-60 sequencing and quantitative PCR. Applied and Environmental Microbiology **71**(2): 867-875.

Hill, J.E., Paccagnella, A., Law, K., Melito, P.L., Woodward, D.L., Price, D.L., Ng, L.-K., Hemmingsen, S.M., and Goh, S.H. 2006a. Identification of *Campylobacter* spp. and discrimination from *Helicobacter* and *Arcobacter* spp. by direct sequencing of PCR-amplified cpn60 sequences and comparison to cpnDB, a chaperonin reference sequence database. Journal of Medical Microbiology **55**(Pt 4): 393-399.

Hill, J.E., Penny, S.L., Crowell, K.G., Goh, S.H., and Hemmingsen, S.M. 2004. cpnDB: a chaperonin sequence database. Genome Research **14**(8): 1669-1675.

Hill, J.E., Seipp, R.P., Betts, M., Hawkins, L., Van Kessel, A.G., Crosby, W.L., and Hemmingsen, S.M. 2002. Extensive profiling of a complex microbial community by high-throughput sequencing. Applied and Environmental Microbiology **68**(6): 3055-3066.

Hill, J.E., Town, J.R., and Hemmingsen, S.M. 2006b. Improved template representation in cpn60 polymerase chain reaction (PCR) product libraries generated from complex templates by application of a specific mixture of PCR primers. Environmental Microbiology **8**(4): 741-746.

Hillier, S.L., Critchlow, C.W., Stevens, C.E., Roberts, M.C., Wolner-Hanssen, P., Eschenbach, D.A., and Holmes, K.K. 1991. Microbiological, epidemiological and clinical correlates of vaginal colonisation by *Mobiluncus* species. Genitourinary Medicine **67**(1): 26-31.

Hillier, S.L., and Lau, R.J. 1997. Vaginal microflora in postmenopausal women who have not received estrogen replacement therapy. Clinical Infectious Diseases **25 Suppl 2**: S123-126.

Hillier, S.L., Nugent, R.P., Eschenbach, D.A., Krohn, M.A., Gibbs, R.S., Martin, D.H., Cotch, M.F., Edelman, R., Pastorek, J.G., 2nd, Rao, A.V., and et al. 1995. Association between bacterial vaginosis and preterm delivery of a low-birth-weight infant. The Vaginal Infections and Prematurity Study Group. New England Journal of Medicine **333**(26): 1737-1742.

Holmes, K.K., Spiegel, C., Amsel, A.R., Eschenbach, D.A., Chen, K.C., and Totten, P. 1981. Nonspecific vaginosis. Scandinavian Journal of Infectious Diseases. Supplementum **26**: 110-114.

Holst, E., Goffeng, A.R., and Andersch, B. 1994. Bacterial vaginosis and vaginal microorganisms in idiopathic premature labor and association with pregnancy outcome. Journal of Clinical Microbiology **32**(1): 176-186.

Holt, J.G. 1994. Bergey's Manual of Determinative Bacteriology. 9 ed. Williams & Wilkins, Baltimore, Md. [u.a.].

Huggins, G.R., and Preti, G. 1976. Volatile constituents of human vaginal secretions. American Journal of Obstetrics and Gynecology **126**(1): 129-136.

Hummelen, R., Fernandes, A.D., Macklaim, J.M., Dickson, R.J., Changalucha, J., Gloor, G.B., and Reid, G. 2010. Deep sequencing of the vaginal microbiota of women with HIV. PLoS ONE **5**(8): e12078.

Hung, W.C., Tseng, S.P., Chen, H.J., Tsai, J.C., Chang, C.H., Lee, T.F., Hsueh, P.R., and Teng, L.J. 2010. Use of groESL as a target for identification of *Abiotrophia, Granulicatella*, and *Gemella* species. Journal of Clinical Microbiology **48**(10): 3532-3538.

Huson, D.H., Richter, D.C., Rausch, C., Dezulian, T., Franz, M., and Rupp, R. 2007. Dendroscope: An interactive viewer for large phylogenetic trees. BMC Bioinformatics **8**: 460.

Ingianni, A., Petruzzelli, S., Morandotti, G., and Pompei, R. 1997. Genotypic differentiation of *Gardnerella vaginalis* by amplified ribosomal DNA restriction analysis (ARDRA). FEMS Immunology and Medical Microbiology **18**(1): 61-66.

Jackman, P.J. 1982. Classification of *Corynebacterium* species from axillary skin by numerical analysis of electrophoretic protein patterns. Journal of Medical Microbiology **15**(4): 485-492.

Jacobsson, B., Pernevi, P., Chidekel, L., and Jorgen Platz-Christensen, J. 2002. Bacterial vaginosis in early pregnancy may predispose for preterm birth and postpartum endometritis. Acta Obstetricia et Gynecologica Scandinavica **81**(11): 1006-1010.

Jakobsson, T., and Forsum, U. 2007. *Lactobacillus iners*: a marker of changes in the vaginal flora? Journal of Clinical Microbiology **45**(9): 3145-3145.

Jarosik, G.P., and Land, C.B. 2000. Identification of a human lactoferrin-binding protein in *Gardnerella vaginalis*. Infection and Immunity **68**(6): 3443-3447.

Jensen, J.S., Borre, M.B., and Dohn, B. 2003. Detection of *Mycoplasma genitalium* by PCR amplification of the 16S rRNA gene. Journal of Clinical Microbiology **41**(1): 261-266.

Jian, W., Zhu, L., and Dong, X. 2001. New approach to phylogenetic analysis of the genus *Bifidobacterium* based on partial HSP60 gene sequences. International Journal of Systematic and Evolutionary Microbiology **51**(Pt 5): 1633-1638.

Joesoef, M.R., Hillier, S.L., Wiknjosastro, G., Sumampouw, H., Linnan, M., Norojono, W., Idajadi, A., and Utomo, B. 1995. Intravaginal clindamycin treatment for bacterial vaginosis: effects on preterm delivery and low birth weight. American Journal of Obstetrics and Gynecology **173**(5): 1527-1531.

Johnson, A.P., and Boustouller, Y.L. 1987. Extra-vaginal infection caused by *Gardnerella vaginalis*. Epidemiology and Infection **98**(2): 131-137.

Jones, B.M., and Al-Mushrif, S. 1997. The determination of phospholipase A2 enzyme activity in the vaginal secretions of pregnant and non-pregnant women with bacterial vaginosis-and in culture exudates of its causative organisms. Journal of Obstetrics and Gynaecology **17**(3): 290-292.

Jones, H.E., Harris, K.A., Azizia, M., Bank, L., Carpenter, B., Hartley, J.C., Klein, N., and Peebles, D. 2009. Differing prevalence and diversity of bacterial species in fetal membranes from very preterm and term labor. PLoS ONE **4**(12): e8205.

Kacerovsky, M., Pavlovsky, M., and Tosner, J. 2009. Preterm premature rupture of the membranes and genital mycoplasmas. Acta Medica (Hradec Kralove) **52**(3): 117-120.

Kampan, N.C., Suffian, S.S., Ithnin, N.S., Muhammad, M., Zakaria, S.Z., and Jamil, M.A. 2011. Evaluation of BV Blue Test Kit for the diagnosis of bacterial vaginosis. Sexual and Reproductive Healthcare **2**(1): 1-5.

Kharsany, A.B., Hoosen, A.A., and Van den Ende, J. 1993. Antimicrobial susceptibilities of *Gardnerella vaginalis*. Antimicrobial Agents and Chemotherapy **37**(12): 2733-2735.

Kim, M.J., Romero, R., Gervasi, M.T., Kim, J.-S., Yoo, W., Lee, D.-C., Mittal, P., Erez, O., Kusanovic, J.P., and Hassan, S.S. 2009. Widespread microbial invasion of the chorioamniotic membranes is a consequence and not a cause of intra-amniotic infection. Laboratory Investigation **89**(8): 924-936.

Koumans, E.H., Markowitz, L.E., and Hogan, V. 2002. Indications for therapy and treatment recommendations for bacterial vaginosis in nonpregnant and pregnant women: a synthesis of data. Clinical Infectious Diseases **35**(Suppl 2): S152-172.

Lamanna, C., and Mallette, M.F. 1954. The cytological basis for the role of the primary dye in the gram stain. Journal of Bacteriology **68**(5): 509-513.

Larsen, B., and Hwang, J. 2010. Mycoplasma, Ureaplasma, and adverse pregnancy outcomes: a fresh look. Infectious Diseases in Obstetrics and Gynecology **2010**.

Leitich, H., Bodner-Adler, B., Brunbauer, M., Kaider, A., Egarter, C., and Husslein, P. 2003. Bacterial vaginosis as a risk factor for preterm delivery: A meta-analysis. American Journal of Obstetrics and Gynecology **189**(1): 139-147.

Leitich, H., and Kiss, H. 2007. Asymptomatic bacterial vaginosis and intermediate flora as risk factors for adverse pregnancy outcome. Bailliere's Best Practice and Research. Clinical Obstetrics and Gynaecology **21**(3): 375-390.

Leopold, S. 1953. Heretofore undescribed organism isolated from the genitourinary system. U S Armed Forces Med J **4**(2): 263-266.

Lewis, W.G., Robinson, L.S., Gilbert, N.M., Perry, J.C., and Lewis, A.L. 2013. Degradation, foraging, and depletion of mucus sialoglycans by the vagina-adapted Actinobacterium *Gardnerella vaginalis*. Journal of Biological Chemistry **288**(17): 12067-12079.

Links, M.G., Chaban, B., Hemmingsen, S.M., Muirhead, K., and Hill, J.E. 2013. mPUMA: a computational approach to microbiota analysis by de novo assembly of operational taxonomic units based on protein-coding barcode sequences. Microbiome **1**(1): 1-7.

Locksmith, G., and Duff, P. 2001. Infection, antibiotics, and preterm delivery. Seminars in Perinatology **25**(5): 295-309.

Macklaim, J.M., Gloor, G.B., Anukam, K.C., Cribby, S., and Reid, G. 2011. At the crossroads of vaginal health and disease, the genome sequence of *Lactobacillus iners*

AB-1. Proceedings of the National Academy of Sciences of the United States of America **108 Suppl 1**: 4688-4695.

Magness, R.R. 1998. Maternal cardiovascular and other physiologic responses to the endocrinology of pregnancy. *In* Endocrinology of pregnancy. Springer. pp. 507-539.

Malone, B.H., Schreiber, M., Schneider, N.J., and Holdeman, L.V. 1975. Obligately anaerobic strains of *Corynebacterium vaginale* (*Haemophilus vaginalis*). Journal of Clinical Microbiology **2**(3): 272-275.

Marconi, C., Donders, G.G., Bellen, G., Brown, D.R., Parada, C.M., and Silva, M.G. 2013. Sialidase activity in aerobic vaginitis is equal to levels during bacterial vaginosis. European Journal of Obstetrics, Gynecology, and Reproductive Biology **167**(2): 205-209.

Marolda, C.L., Vicarioli, J., and Valvano, M.A. 2004. Wzx proteins involved in biosynthesis of O antigen function in association with the first sugar of the O-specific lipopolysaccharide subunit. Microbiology **150**(Pt 12): 4095-4105.

Martin, D.H., Zozaya, M., Lillis, R., Miller, J., and Ferris, M.J. 2012. The microbiota of the human genitourinary tract: trying to see the forest through the trees. Transactions of the American Clinical and Climatological Association **123**: 242-256.

Martín-Platero, A., Valdivia, E., Maqueda, M., and Martínez-Bueno, M. 2007. Fast, convenient, and economical method for isolating genomic DNA from lactic acid bacteria using a modification of the protein "salting-out" procedure. Analytical Biochemistry **366**(1): 102-104.

Martius, J., Krohn, M.A., Hillier, S.L., Stamm, W.E., Holmes, K.K., and Eschenbach, D.A. 1988. Relationships of vaginal *Lactobacillus* species, cervical *Chlamydia trachomatis*, and bacterial vaginosis to preterm birth. Obstetrics and Gynecology **71**(1): 89-95.

Mascagni, J.R., and Miller, L.H. 2004. A descriptive correlational study of bacterial vaginosis in pregnancy and its association with preterm birth: implications for advanced practice nurses. Journal of the American Academy of Nurse Practitioners **16**(12): 555-560.

Mass, S.B., Brennan, J.P., Silverman, N., and van Hoeven, K.H. 1999. Association between a shift in vaginal flora on Papanicolaou smear and acute chorioamnionitis and preterm delivery. Diagnostic Cytopathology **21**(1): 7-9.

Mastrobattista, J.M., Bishop, K.D., and Newton, E.R. 2000. Wet smear compared with gram stain diagnosis of bacterial vaginosis in asymptomatic pregnant women. Obstetrics and Gynecology **96**(4): 504-506.

McGregor, J.A., French, J.I., Jones, W., Milligan, K., McKinney, P.J., Patterson, E., and Parker, R. 1994. Bacterial vaginosis is associated with prematurity and vaginal fluid mucinase and sialidase: results of a controlled trial of topical clindamycin cream.

American Journal of Obstetrics and Gynecology **170**(4): 1048-1059; discussion 1059-1060.

McLean, N.W., and McGroarty, J.A. 1996. Growth inhibition of metronidazole-susceptible and metronidazole-resistant strains of *Gardnerella vaginalis* by Lactobacilli in vitro. Applied and Environmental Microbiology **62**(3): 1089-1092.

Menard, J.P., Fenollar, F., Henry, M., Bretelle, F., and Raoult, D. 2008. Molecular quantification of *Gardnerella vaginalis* and *Atopobium vaginae* loads to predict bacterial vaginosis. Clinical Infectious Diseases **47**(1): 33-43.

Menard, J.P., Mazouni, C., Salem-Cherif, I., Fenollar, F., Raoult, D., Boubli, L., Gamerre, M., and Bretelle, F. 2010. High vaginal concentrations of *Atopobium vaginae* and *Gardnerella vaginalis* in women undergoing preterm labor. Obstetrics and Gynecology **115**(1): 134-140.

Mercer, B.M. 2003. Preterm premature rupture of the membranes. Obstetrics and Gynecology **101**(1): 178-193.

Mikamo, H., Sato, Y., Hayasaki, Y., Kawazoe, K., Izumi, K., Ito, K., Yamada, Y., and Tamaya, T. 1999. Intravaginal bacterial flora in patients with uterine cervical cancer. High incidence of detection of *Gardnerella vaginalis*. J Infect Chemother **5**(2): 82-85.

Minana-Galbis, D., Urbizu-Serrano, A., Farfan, M., Fuste, M.C., and Loren, J.G. 2009. Phylogenetic analysis and identification of *Aeromonas* species based on sequencing of the cpn60 universal target. International Journal of Systematic and Evolutionary Microbiology **59**(8): 1976-1983.

Moncla, B.J., and Braham, P. 1989. Detection of sialidase (neuraminidase) activity in Actinomyces species by using 2'-(4-methylumbelliferyl)alpha-D-N-acetylneuraminic acid in a filter paper spot test. Journal of Clinical Microbiology **27**(1): 182-184.

Moxon, E., and Kroll, J. 1990. The role of bacterial polysaccharide capsules as virulence factors. *In* Bacterial capsules. Springer. pp. 65-85.

Myer, L., Kuhn, L., Stein, Z.A., Wright, T.C., Jr., and Denny, L. 2005. Intravaginal practices, bacterial vaginosis, and women's susceptibility to HIV infection: epidemiological evidence and biological mechanisms. Lancet Infectious Diseases **5**(12): 786-794.

Nauth, H.F., and Haas, M. 1985. Cytologic and histologic observations on the sex hormone dependence of the vulva. Journal of Reproductive Immunology **30**(9): 667-674.

Nelson, D.B., Hanlon, A., Nachamkin, I., Haggerty, C., Mastrogiannis, D.S., Liu, C., and Fredricks, D.N. 2014. Early pregnancy changes in bacterial vaginosis-associated bacteria and preterm delivery. Paediatric and Perinatal Epidemiology **28**(2): 88-96.

Ness, R.B., Hillier, S.L., Kip, K.E., Soper, D.E., Stamm, C.A., McGregor, J.A., Bass, D.C., Sweet, R.L., Rice, P., and Richter, H.E. 2004. Bacterial vaginosis and risk of pelvic inflammatory disease. Obstetrics and Gynecology **104**(4): 761-769.

Nugent, R.P., Krohn, M.A., and Hillier, S.L. 1991. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. Journal of Clinical Microbiology **29**(2): 297-301.

Numanovic, F., Hukic, M., Nurkic, M., Gegic, M., Delibegovic, Z., Imamovic, A., and Pasic, S. 2008. Importance of isolation and biotypization of *Gardnerella vaginalis* in diagnosis of bacterial vaginosis. Bosnian Journal of Basic Medical Sciences **8**(3): 270-276.

O'Hanlon, D.E., Moench, T.R., and Cone, R.A. 2013. Vaginal pH and microbicidal lactic acid when lactobacilli dominate the microbiota. PLoS ONE **8**(11): e80074.

Oakley, B.B., Fiedler, T.L., Marrazzo, J.M., and Fredricks, D.N. 2008. Diversity of human vaginal bacterial communities and associations with clinically defined bacterial vaginosis. Applied and Environmental Microbiology **74**(15): 4898-4909.

Oksnen, J., Blanchet, F., Kindt, R., Legendre, P., and Minchin, P. 2012. vegan: Community Ecology Package. R Package V. 2.0.

Oleen-Burkey, M.A., and Hillier, S.L. 1995. Pregnancy complications associated with bacterial vaginosis and their estimated costs. Infectious Diseases in Obstetrics and Gynecology **3**(4): 149-157.

Onderdonk, A.B., Zamarchi, G.R., Walsh, J.A., Mellor, R.D., Munoz, A., and Kass, E.H. 1986. Methods for quantitative and qualitative evaluation of vaginal microflora during menstruation. Applied and Environmental Microbiology **51**(2): 333-339.

Overduin, P., Boos, W., and Tommassen, J. 1988. Nucleotide sequence of the ugp genes of *Escherichia coli* K-12: homology to the maltose system. Molecular Microbiology **2**(6): 767-775.

Pandit, L., and Ouslander, J.G. 1997. Postmenopausal vaginal atrophy and atrophic vaginitis. American Journal of the Medical Sciences **314**(4): 228-231.

Paramel Jayaprakash, T., Schellenberg, J.J., and Hill, J.E. 2012. Resolution and characterization of distinct cpn60-based subgroups of *Gardnerella vaginalis* in the vaginal microbiota. PLoS ONE **7**(8): e43009.

Patterson, J.L., Stull-Lane, A., Girerd, P.H., and Jefferson, K.K. 2010. Analysis of adherence, biofilm formation and cytotoxicity suggests a greater virulence potential of *Gardnerella vaginalis* relative to other bacterial-vaginosis-associated anaerobes. Microbiology **156**(Pt 2): 392-399.

Paul, V.K., Gupta, U., Singh, M., Nag, V.L., Takkar, D., and Bhan, M.K. 1998. Association of genital mycoplasma colonization with low birth weight. International Journal of Gynaecology and Obstetrics **63**(2): 109-114.

Pediatrics, A.A.o. 2006. The APGAR score. Advances in Neonatal Care **6**(4): 220-223.

Pepin, J., Labbe, A.C., Khonde, N., Deslandes, S., Alary, M., Dzokoto, A., Asamoah-Adu, C., Meda, H., and Frost, E. 2005. *Mycoplasma genitalium*: an organism commonly associated with cervicitis among west African sex workers. Sexually Transmitted Infections **81**(1): 67-72.

Petkau, A., Stuart-Edwards, M., Stothard, P., and Van Domselaar, G. 2010. Interactive microbial genome visualization with GView. Bioinformatics **26**(24): 3125-3126.

Petricevic, L., Domig, K.J., Nierscher, F.J., Sandhofer, M.J., Fidesser, M., Krondorfer, I., Husslein, P., Kneifel, W., and Kiss, H. 2014. Characterisation of the vaginal *Lactobacillus* microbiota associated with preterm delivery. Scientific Reports **4**: 5136.

Petricevic, L., Domig, K.J., Nierscher, F.J., Sandhofer, M.J., Krondorfer, I., Kneifel, W., and Kiss, H. 2013. Differences in the vaginal lactobacilli of postmenopausal women and influence of rectal lactobacilli. Climacteric **16**(3): 356-361.

Piot, P., van Dyck, E., Goodfellow, M., and Falkow, S. 1980. A taxonomic study of *Gardnerella vaginalis* (*Haemophilus vaginalis*) Gardner and Dukes 1955. Journal of General Microbiology **119**(2): 373-396.

Piot, P., Van Dyck, E., Peeters, M., Hale, J., Totten, P.A., and Holmes, K.K. 1984. Biotypes of *Gardnerella vaginalis*. Journal of Clinical Microbiology **20**(4): 677-679.

Pleckaityte, M., Janulaitiene, M., Lasickiene, R., and Zvirbliene, A. 2012a. Genetic and biochemical diversity of *Gardnerella vaginalis* strains isolated from women with bacterial vaginosis. FEMS Immunology and Medical Microbiology **65**(1): 69-77.

Pleckaityte, M., Zilnyte, M., and Zvirbliene, A. 2012b. Insights into the CRISPR/Cas system of *Gardnerella vaginalis*. BMC Microbiology **12**: 301.

Powell, K. 2013. Bacterial vaginosis: risks and treatment. British Journal of Nursing **22**(18): 1044.

Pybus, V., and Onderdonk, A.B. 1997. Evidence for a commensal, symbiotic relationship between *Gardnerella vaginalis* and *Prevotella bivia* involving ammonia: potential significance for bacterial vaginosis. Journal of Infectious Diseases **175**(2): 406-413.

Pybus, V., and Onderdonk, A.B. 1998. A commensal symbiosis between *Prevotella bivia* and *Peptostreptococcus anaerobius* involves amino acids: potential significance to the pathogenesis of bacterial vaginosis. FEMS Immunology and Medical Microbiology **22**(4): 317-327.

Queiros da Mota, V., Prodhom, G., Yan, P., Hohlfheld, P., Greub, G., and Rouleau, C. 2013. Correlation between placental bacterial culture results and histological chorioamnionitis: a prospective study on 376 placentas. Journal of Clinical Pathology **66**(3): 243-248.

Ravel, J., Gajer, P., Abdo, Z., Schneider, G.M., Koenig, S.S., McCulle, S.L., Karlebach, S., Gorle, R., Russell, J., Tacket, C.O., Brotman, R.M., Davis, C.C., Ault, K., Peralta, L., and Forney, L.J. 2011. Vaginal microbiome of reproductive-age women. Proceedings of the National Academy of Sciences of the United States of America **108 Suppl 1**: 4680-4687.

Reeder, J., and Knight, R. 2009. The 'rare biosphere': a reality check. Nature Methods **6**(9): 636-637.

Reyn, A., Birch-Andersen, A., and Lapage, S.P. 1966. An electron microscope study of thin sections of *Haemophilus vaginalis* (Gardner and Dukes) and some possibly related species. Canadian Journal of Microbiology **12**(6): 1125-1136.

Rice, P., Longden, I., and Bleasby, A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends in Genetics **16**(6): 276-277.

Richter, M., and Rossello-Mora, R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proceedings of the National Academy of Sciences of the United States of America **106**(45): 19126-19131.

Roberton, A.M., Wiggins, R., Horner, P.J., Greenwood, R., Crowley, T., Fernandes, A., Berry, M., and Corfield, A.P. 2005. Novel bacterial mucinase, glycosulfatase, is associated with bacterial vaginosis. Journal of Clinical Microbiology **43**(11): 5504-5508.

Romero, R., Hassan, S.S., Gajer, P., Tarca, A.L., Fadrosh, D.W., Bieda, J., Chaemsaithong, P., Miranda, J., Chaiworapongsa, T., and Ravel, J. 2014a. The vaginal microbiota of pregnant women who subsequently have spontaneous preterm labor and delivery and those with a normal delivery at term. Microbiome **2**(1): 18.

Romero, R., Hassan, S.S., Gajer, P., Tarca, A.L., Fadrosh, D.W., Nikita, L., Galuppi, M., Lamont, R.F., Chaemsaithong, P., Miranda, J., Chaiworapongsa, T., and Ravel, J. 2014b. The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. Microbiome **2**(1): 4.

Sadhu, K., Domingue, P.A., Chow, A.W., Nelligan, J., Cheng, N., and Costerton, J.W. 1989. *Gardnerella vaginalis* has a gram-positive cell-wall ultrastructure and lacks classical cell-wall lipopolysaccharide. Journal of Medical Microbiology **29**(3): 229-235.

Saitou, N., and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution **4**(4): 406-425.

Sakamoto, M., and Ohkuma, M. 2010. Usefulness of the hsp60 gene for the identification and classification of Gram-negative anaerobic rods. Journal of Medical Microbiology **59**(Pt 11): 1293-1302.

Sakamoto, M., Suzuki, N., and Benno, Y. 2010. hsp60 and 16S rRNA gene sequence relationships among species of the genus *Bacteroides* with the finding that *Bacteroides suis* and *Bacteroides tectus* are heterotypic synonyms of *Bacteroides pyogenes*. International Journal of Systematic and Evolutionary Microbiology **60**(Pt 12): 2984-2990.

Santiago, G.L., Cools, P., Verstraelen, H., Trog, M., Missine, G., El Aila, N., Verhelst, R., Tency, I., Claeys, G., Temmerman, M., and Vaneechoutte, M. 2011a. Longitudinal study of the dynamics of vaginal microflora during two consecutive menstrual cycles. PLoS ONE **6**(11): e28180.

Santiago, G.L., Deschaght, P., El Aila, N., Kiama, T.N., Verstraelen, H., Jefferson, K.K., Temmerman, M., and Vaneechoutte, M. 2011b. *Gardnerella vaginalis* comprises three distinct genotypes of which only two produce sialidase. American Journal of Obstetrics and Gynecology **204**(5): 450 e451-457.

Santiago, G.L., Tency, I., Verstraelen, H., Verhelst, R., Trog, M., Temmerman, M., Vancoillie, L., Decat, E., Cools, P., and Vaneechoutte, M. 2012. Longitudinal qPCR study of the dynamics of *L. crispatus*, *L. iners*, *A. vaginae,* (sialidase positive) *G. vaginalis*, and *P. bivia* in the vagina. PLoS ONE **7**(9): e45281.

Schellenberg, J., Links, M.G., Hill, J.E., Dumonceaux, T.J., Peters, G.A., Tyler, S., Ball, T.B., Severini, A., and Plummer, F.A. 2009a. Pyrosequencing of the chaperonin-60 universal target as a tool for determining microbial community composition. Applied and Environmental Microbiology **75**(9): 2889-2898.

Schellenberg, J., Links, M.G., Hill, J.E., Hemmingsen, S.M., Peters, G.A., and Dumonceaux, T.J. 2011a. Pyrosequencing of chaperonin-60 (cpn60) amplicons as a means of determining microbial community composition. Methods in Molecular Biology **733**: 143-158.

Schellenberg, J.J., Ball, T.B., Kimani, J., Jaoko, W., Wachihi, C., Njeri Mungai, J., Dumonceaux, T.J., Hill, J.E., Links, M.G., and Plummer, F.A. 2009b. Defining bacterial vaginosis through in-depth culture-based and molecular characterization of the vaginal microbiome in exposed seronegative commercial sex workers from Nairobi, Kenya. *In* International Symposium on Natural Immunity to HIV., Winnipeg, MB

Schellenberg, J.J., Dumonceaux, T.J., Hill, J.E., Kimani, J., Jaoko, W., Wachihi, C., Mungai, J.N., Lane, M., Fowke, K.R., Ball, T.B., and Plummer, F.A. 2012. Selection, phenotyping and identification of acid and hydrogen peroxide producing bacteria from vaginal samples of Canadian and East African women. PLoS ONE **7**(7): e41217.

Schellenberg, J.J., Links, M.G., Hill, J.E., Dumonceaux, T.J., Kimani, J., Jaoko, W., Wachihi, C., Mungai, J.N., Peters, G.A., Tyler, S., Graham, M., Severini, A., Fowke,

K.R., Ball, T.B., and Plummer, F.A. 2011b. Molecular definition of vaginal microbiota in East African commercial sex workers. Applied and Environmental Microbiology.

Schlicht, M.J., Lovrich, S.D., Sartin, J.S., Karpinsky, P., Callister, S.M., and Agger, W.A. 2004. High prevalence of genital mycoplasmas among sexually active young adults with urethritis or cervicitis symptoms in La Crosse, Wisconsin. Journal of Clinical Microbiology **42**(10): 4636-4640.

Schoonmaker, J.N., Lunt, B.D., Lawellin, D.W., French, J.I., Hillier, S.L., and McGregor, J.A. 1991. A new proline aminopeptidase assay for diagnosis of bacterial vaginosis. American Journal of Obstetrics and Gynecology **165**(3): 737-742.

Shanker, S., Toohey, M., and Munro, R. 1982. In vitro activity of seventeen antimicrobial agents against *Gardnerella vaginalis*. European Journal of Clinical Microbiology **1**(5): 298-300.

Shipitsyna, E., Roos, A., Datcu, R., Hallen, A., Fredlund, H., Jensen, J.S., Engstrand, L., and Unemo, M. 2013. Composition of the vaginal microbiota in women of reproductive age–sensitive and specific molecular diagnosis of bacterial vaginosis is possible? PLoS ONE **8**(4): e60670.

Shujatullah, F., Khan, H.M., Khatoon, R., Rabbani, T., and Malik, A. 2010. An evaluation of OSOM BV blue test in the diagnosis of bacterial vaginosis. Asian Pacific Journal of Tropical Medicine **3**(7): 574-576.

Simhan, H.N., and Canavan, T.P. 2005. Preterm premature rupture of membranes: diagnosis, evaluation and management strategies. BJOG **112 Suppl 1**: 32-37.

Simoes, J.A., Discacciati, M.G., Brolazo, E.M., Portugal, P.M., Dini, D.V., and Dantas, M.C. 2006. Clinical diagnosis of bacterial vaginosis. International Journal of Gynaecology and Obstetrics **94**(1): 28-32.

Sjoberg, I., Cajander, S., and Rylander, E. 1988. Morphometric characteristics of the vaginal epithelium during the menstrual cycle. Gynecologic and Obstetric Investigation **26**(2): 136-144.

Sobel, J.D. 2000. Bacterial vaginosis. Annual Review of Medicine **51**: 349-356.

Soong, D., and Einarson, A. 2009. Vaginal yeast infections during pregnancy. Canadian Family Physician **55**(3): 255-256.

Soper, D.E., Brockwell, N.J., Dalton, H.P., and Johnson, D. 1994. Observations concerning the microbial etiology of acute salpingitis. American Journal of Obstetrics and Gynecology **170**(4): 1008-1014; discussion 1014-1007.

Spear, G.T., French, A.L., Gilbert, D., Zariffard, M.R., Mirmonsef, P., Sullivan, T.H., Spear, W.W., Landay, A., Micci, S., Lee, B.H., and Hamaker, B.R. 2014. Human alpha-

amylase Present in Lower-Genital-Tract Mucosal Fluid Processes Glycogen to Support Vaginal Colonization by Lactobacillus. Journal of Infectious Diseases.

Spiegel, C.A. 1991. Bacterial vaginosis. Clinical Microbiology Reviews **4**(4): 485-502.

Srinivasan, S., Hoffman, N.G., Morgan, M.T., Matsen, F.A., Fiedler, T.L., Hall, R.W., Ross, F.J., McCoy, C.O., Bumgarner, R., Marrazzo, J.M., and Fredricks, D.N. 2012. Bacterial communities in women with bacterial vaginosis: high resolution phylogenetic analyses reveal relationships of microbiota to clinical criteria. PLoS ONE **7**(6): e37818.

Srinivasan, S., Morgan, M.T., Liu, C., Matsen, F.A., Hoffman, N.G., Fiedler, T.L., Agnew, K.J., Marrazzo, J.M., and Fredricks, D.N. 2013. More than meets the eye: associations of vaginal bacteria with gram stain morphotypes using molecular phylogenetic analysis. PLoS ONE **8**(10): e78633.

Staden, R., Beal, K.F., and Bonfield, J.K. 2000. The Staden package, 1998. Methods in Molecular Biology **132**: 115-130.

Stanley, J., Linton, D., Burnens, A.P., Dewhirst, F.E., Owen, R.J., Porter, A., On, S.L., and Costas, M. 1993. *Helicobacter canis* sp. nov., a new species from dogs: an integrated study of phenotype and genotype. Journal of General Microbiology **139**(10): 2495-2504.

Stewart-Tull, D. 1964. Evidence that vaginal Lactobacilli do not ferment glycogen. American Journal of Obstetrics and Gynecology **88**: 676.

Summers, P.R. 2010. Vaginal secretions with age. *In* Textbook of Aging Skin. Springer. pp. 247-251.

Swidsinski, A., Doerffel, Y., Loening-Baucke, V., Swidsinski, S., Verstraelen, H., Vaneechoutte, M., Lemm, V., Schilling, J., and Mendling, W. 2010. *Gardnerella* biofilm involves females and males and is transmitted sexually. Gynecologic and Obstetric Investigation **70**(4): 256-263.

Swidsinski, A., Mendling, W., Loening-Baucke, V., Ladhoff, A., Swidsinski, S., Hale, L.P., and Lochs, H. 2005. Adherent biofilms in bacterial vaginosis. Obstetrics and Gynecology **106**(5 Pt 1): 1013-1023.

Swidsinski, A., Mendling, W., Loening-Baucke, V., Swidsinski, S., Dorffel, Y., Scholze, J., Lochs, H., and Verstraelen, H. 2008. An adherent *Gardnerella vaginalis* biofilm persists on the vaginal epithelium after standard therapy with oral metronidazole. American Journal of Obstetrics and Gynecology **198**(1): 97 e91-96.

Taylor-Robinson, D. 1996. Infections due to species of *Mycoplasma* and *Ureaplasma*: an update. Clinical Infectious Diseases **23**(4): 671-682; quiz 683-674.

Taylor-Robinson, D., and Lamont, R.F. 2011. Mycoplasmas in pregnancy. BJOG **118**(2): 164-174.

Teixeira, G.S., Soares-Brandao, K.L., Branco, K.M., Sampaio, J.L., Nardi, R.M., Mendonca, M., Almeida, R.B., Farias, L.M., Carvalho, M.A., and Nicoli, J.R. 2010. Antagonism and synergism in *Gardnerella vaginalis* strains isolated from women with bacterial vaginosis. Journal of Medical Microbiology **59**(Pt 8): 891-897.

Teng, L.J., Ho, S.W., Ho, H.N., Liaw, S.J., Lai, H.C., and Luh, K.T. 1995. Rapid detection and biovar differentiation of *Ureaplasma urealyticum* in clinical specimens by PCR. Journal of the Formosan Medical Association **94**(7): 396-400.

Teng, L.J., Zheng, X., Glass, J.I., Watson, H.L., Tsai, J., and Cassell, G.H. 1994. *Ureaplasma urealyticum* biovar specificity and diversity are encoded in multiple-banded antigen gene. Journal of Clinical Microbiology **32**(6): 1464-1469.

Thies, F.L., Konig, W., and Konig, B. 2007. Rapid characterization of the normal and disturbed vaginal microbiota by application of 16S rRNA gene terminal RFLP fingerprinting. Journal of Medical Microbiology **56**(Pt 6): 755-761.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research **22**(22): 4673-4680.

Tomusiak, A., Strus, M., and Heczko, P.B. 2011. Antibiotic resistance of *Gardnerella vaginalis* isolated from cases of bacterial vaginosis. Ginekologia Polska **82**(12): 900-904.

Tosun, I., Alpay Karaoglu, S., Ciftci, H., Buruk, C.K., Aydin, F., Kilic, A.O., and Erturk, M. 2007. Biotypes and antibiotic resistance patterns of *Gardnerella vaginalis* strains isolated from healthy women and women with bacterial vaginosis. Mikrobiyoloji Bulteni **41**(1): 21-27.

Turovskiy, Y., Sutyak Noll, K., and Chikindas, M.L. 2011. The aetiology of bacterial vaginosis. Journal of Applied Microbiology **110**(5): 1105-1128.

Van Dyck, E., Ieven, M., Pattyn, S., Van Damme, L., and Laga, M. 2001. Detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* by enzyme immunoassay, culture, and three nucleic acid amplification tests. Journal of Clinical Microbiology **39**(5): 1751-1756.

van Kuppeveld, F.J., van der Logt, J.T., Angulo, A.F., van Zoest, M.J., Quint, W.G., Niesters, H.G., Galama, J.M., and Melchers, W.J. 1992. Genus- and species-specific identification of mycoplasmas by 16S rRNA amplification. Applied and Environmental Microbiology **58**(8): 2606-2615.

Vanhoof, G., Goossens, F., De Meester, I., Hendriks, D., and Scharpe, S. 1995. Proline motifs in peptides and their biological processing. FASEB Journal **9**(9): 736-744.

Vasquez, A., Jakobsson, T., Ahrne, S., Forsum, U., and Molin, G. 2002. Vaginal lactobacillus flora of healthy Swedish women. Journal of Clinical Microbiology **40**(8): 2746-2749.

Verbeke, T.J., Sparling, R., Hill, J.E., Links, M.G., Levin, D., and Dumonceaux, T.J. 2011. Predicting relatedness of bacterial genomes using the chaperonin-60 universal target (cpn60 UT): application to *Thermoanaerobacter* species. . Systematic and Applied Microbiology **34**(3): 171-179.

Verhelst, R., Verstraelen, H., Claeys, G., Verschraegen, G., Delanghe, J., Van Simaey, L., De Ganck, C., Temmerman, M., and Vaneechoutte, M. 2004. Cloning of 16S rRNA genes amplified from normal and disturbed vaginal microflora suggests a strong association between *Atopobium vaginae*, *Gardnerella vaginalis* and bacterial vaginosis. BMC Microbiology **4**: 16.

Verma, A., Okun, N.B., Maguire, T.O., and Mitchell, B.F. 1999. Morbidity assessment index for newborns: a composite tool for measuring newborn health. American Journal of Obstetrics and Gynecology **181**(3): 701-708.

Vermette, C.J., Russell, A.H., Desai, A.R., and Hill, J.E. 2010. Resolution of phenotypically distinct strains of *Enterococcus* spp. in a complex microbial community using cpn60 universal target sequencing. Microbial Ecology **59**(1): 14-24.

Verstraelen, H., Verhelst, R., Claeys, G., De Backer, E., Temmerman, M., and Vaneechoutte, M. 2009. Longitudinal analysis of the vaginal microflora in pregnancy suggests that *L. crispatus* promotes the stability of the normal vaginal microflora and that *L. gasseri* and/or *L. iners* are more conducive to the occurrence of abnormal vaginal microflora. BMC Microbiology **9**(1): 116.

Waites, K.B., Katz, B., and Schelonka, R.L. 2005. Mycoplasmas and ureaplasmas as neonatal pathogens. Clinical Microbiology Reviews **18**(4): 757-789.

Wallenburg, H.C., and Wladimiroff, J.W. 1976. Recurrence of vulvovaginal candidosis during pregnancy. Comparison of miconazole vs nystatin treatment. Obstetrics and Gynecology **48**(4): 491-494.

Wang, X., Buhimschi, C.S., Temoin, S., Bhandari, V., Han, Y.W., and Buhimschi, I.A. 2013. Comparative microbial analysis of paired amniotic fluid and cord blood from pregnancies complicated by preterm birth and early-onset neonatal sepsis. PLoS ONE **8**(2): e56131.

Wasiela, M., Krzeminski, Z., Hanke, W., and Kalinka, J. 2003. Association between genital mycoplasmas and risk of preterm delivery. Medycyna Wieku Rozwojowego **7**(3 Suppl 1): 211-216.

Welch, R.A. 1991. Pore-forming cytolysins of gram-negative bacteria. Molecular Microbiology **5**(3): 521-528.

Wen, A., Srinivasan, U., Goldberg, D., Owen, J., Marrs, C.F., Misra, D., Wing, D.A., Ponnaluri, S., Miles-Jay, A., Bucholz, B., Abbas, K., and Foxman, B. 2014. Selected vaginal bacteria and risk of preterm birth: an ecological perspective. Journal of Infectious Diseases **209**(7): 1087-1094.

Wiggins, R., Hicks, S.J., Soothill, P.W., Millar, M.R., and Corfield, A.P. 2001. Mucinases and sialidases: their role in the pathogenesis of sexually transmitted infections in the female genital tract. Sexually Transmitted Infections **77**(6): 402-408.

Wolrath, H., Forsum, U., Larsson, P.G., and Boren, H. 2001. Analysis of bacterial vaginosis-related amines in vaginal fluid by gas chromatography and mass spectrometry. Journal of Clinical Microbiology **39**(11): 4026-4031.

Wylie, J.G., and Henderson, A. 1969. Identity and glycogen-fermenting ability of lactobacilli isolated from the vagina of pregnant women. Journal of Medical Microbiology **2**(3): 363-366.

Yeoman, C.J., Yildirim, S., Thomas, S.M., Durkin, A.S., Torralba, M., Sutton, G., Buhay, C.J., Ding, Y., Dugan-Rocha, S.P., Muzny, D.M., Qin, X., Gibbs, R.A., Leigh, S.R., Stumpf, R., White, B.A., Highlander, S.K., Nelson, K.E., and Wilson, B.A. 2010. Comparative genomics of *Gardnerella vaginalis* strains reveals substantial differences in metabolic and virulence potential. PLoS ONE **5**(8): e12411.

Yoshida, T., Maeda, S., Deguchi, T., and Ishiko, H. 2002. Phylogeny-based rapid identification of mycoplasmas and ureaplasmas from urethritis patients. Journal of Clinical Microbiology **40**(1): 105-110.

Yudin, M.H., van Schalkwyk, J., Van Eyk, N., Boucher, M., Castillo, E., Cormier, B., Gruslin, A., Money, D.M., Murphy, K., Ogilvie, G., Paquet, C., Steenbeek, A., Wong, T., Gagnon, R., Hudon, L., Basso, M., Bos, H., Delisle, M.F., Farine, D., Grabowska, K., Menticoglou, S., Mundle, W.R., Murphy-Kaulbeck, L.C., Ouellet, A., Pressey, T., and Roggensack, A. 2009. Antibiotic therapy in preterm premature rupture of the membranes. Journal of Obstetrics and Gynaecology Canada **31**(9): 863-867, 868-874.

Zaragoza, O., Chrisman, C.J., Castelli, M.V., Frases, S., Cuenca-Estrella, M., Rodriguez-Tudela, J.L., and Casadevall, A. 2008. Capsule enlargement in *Cryptococcus neoformans* confers resistance to oxidative stress suggesting a mechanism for intracellular survival. Cellular Microbiology **10**(10): 2043-2057.

Zaragoza, O., Rodrigues, M.L., De Jesus, M., Frases, S., Dadachova, E., and Casadevall, A. 2009. The capsule of the fungal pathogen *Cryptococcus neoformans*. Advances in Applied Microbiology **68**: 133-216.

Zariffard, M.R., Saifuddin, M., Sha, B.E., and Spear, G.T. 2002. Detection of bacterial vaginosis-related organisms by real-time PCR for Lactobacilli, *Gardnerella vaginalis* and *Mycoplasma hominis*. FEMS Immunology and Medical Microbiology **34**(4): 277-281.

Zeigler, D.R. 2003. Gene sequences useful for predicting relatedness of whole genomes in bacteria. International Journal of Systematic and Evolutionary Microbiology **53**(Pt 6): 1893-1900.

Zhang, X., Xu, X., Li, J., Li, N., Yan, T., and Ju, X. 2002. Relationship between vaginal sialidase bacteria vaginosis and chorioammionitis. Zhonghua Fu Chan Ke Za Zhi (Chinese Journal of Obstetrics and Gynecology) **37**(10): 588-590.

Zhou, X., Bent, S.J., Schneider, M.G., Davis, C.C., Islam, M.R., and Forney, L.J. 2004. Characterization of vaginal microbial communities in adult healthy women using cultivation-independent methods. Microbiology-Sgm **150**: 2565-2573.

Zinnemann, K., and Turner, G.C. 1963. The taxonomic position of "*Haemophilus vaginalis*" [*Corynebacterium vaginale*]. The Journal of Pathology and Bacteriology **85**(1): 213-219.

Zozaya-Hinchliffe, M., Martin, D.H., and Ferris, M.J. 2008. Prevalence and abundance of uncultivated *Megasphaera*-like bacteria in the human vaginal environment. Applied and Environmental Microbiology **74**(5): 1656-1659.

Zuppardo, A.B., and Siebeling, R.J. 1998. An epimerase gene essential for capsule synthesis in *Vibrio vulnificus*. Infection and Immunity **66**(6): 2601-2606.