CHARACTERIZATION OF THE VAGINAL MICROBIOME IN PREGNANCY

A Thesis Submitted to the College of Graduate and Postdoctoral Studies

In Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy

In the Department of Veterinary Microbiology

University of Saskatchewan

Saskatoon

By

ALINE COSTA DE FREITAS

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

Requests for permission to copy or to make other uses of materials in this thesis in whole or part should be addressed to:

Head of Department of Veterinary Microbiology
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5B4
Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9
Canada

ABSTRACT

The vaginal microbiome plays an important role in women's reproductive health. Imbalances in this microbiota are associated with bacterial vaginosis, increased susceptibility to sexually transmitted infections, and negative reproductive outcomes. The causes of such variations, however, are poorly understood. A healthy vaginal microbiota is defined as *Lactobacillus*-dominated and an overgrowth of other species is often associated with unhealthy conditions. An appreciation of "atypical" microbiomes in healthy women, such as *Bifidobacterium*-dominated, has been gradually increasing. Although bifidobacteria play an important role in gut health, vaginal bifidobacteria have not yet been fully characterized.

In this study, a baseline description of the "healthy" vaginal microbiome in pregnancy has been established based on cpn60 gene amplicon sequencing. The vaginal microbiota of pregnant women relative to non-pregnant women had lower richness and diversity, higher *Lactobacillus* abundance and lower Mollicutes/*Ureaplasma* prevalence. This gives a better understanding of the vaginal microbiome in healthy pregnancies and provides a control group for a subsequent comparison to women who experienced preterm birth. An association between Mollicutes and preterm was confirmed, and further suggested that a more rich and diverse microbiome is associated with prematurity. To better understand the relationship between reproductive outcomes and microbiota, an improved definition of the healthy microbiome is also needed, which should include evaluation of "atypical" microbiomes, such as *Bifidobacterium*-dominated. Phenotypic characterization of vaginal bifidobacteria indicated that they have health promoting characteristics similar to beneficial vaginal lactobacilli. Considering the importance of bifidobacteria as one of the primary colonizers of the neonatal gut, the genomes of vaginal and gut isolates of *Bifidobacterium breve* and *Bifidobacterium longum* were compared. Results indicated that vaginal and gut microbiomes are colonized by a shared community of *Bifidobacterium*, which may be transferred from mother to infant.

Taken together, the results presented in this thesis provide a better understanding of the vaginal microbiome of pregnant women with low and high risk for preterm birth. It also improves the understanding of a healthy microbiome by phenotypically characterizing vaginal bifidobacteria, and contributes to elucidate aspects of bifidobacteria ecology by comparing the genomes of vaginal and gut bifidobacteria.

ACKNOWLEDGEMENTS

This thesis project could only be completed with the collaboration and support of many people to whom I am very grateful.

First and foremost, I would like to thank my supervisor Dr. Janet Hill for giving me the opportunity to be part of the Hill lab. I am grateful for her guidance, patience, kindness, ability to transform any moment in a teaching moment, encouragement in moments of uncertainty, replying/reviewing materials in a record time, teaching me so much about Science and pushing me beyond my comfort zone (which today I really appreciate!). Thank you!

I also would like to thank the members of my Research Advisory Committee – Dr. Joseph Rubin, Dr. Aaron White, Dr. Jo-Anne Dillon and Dr. Vikram Misra – for their insightful comments that undoubtedly improved the quality of this work. To the VOGUE team, in special Dr. Bocking and Dr. Money, thank you for providing me the opportunity to participate in this interdisciplinary group which immensely contributed to my professional development.

Thanks to the Hill lab members/alumni and aggregates: Bonnie, John, Matthew, Teenus, Isha, Matheus, Jason, Lisa, Mo, Niradha, Salahuddin, Maodong, Shakya, Pashupati, Sonu, Arinjay, Roman, Carolina, Lana, Linda and GMP members. Every seminar, lab meeting, "office" talks and chit chatting contributed for my professional and personal growth, and made everything more enjoyable and fun! A special thanks to Champika Fernando who not only taught me so much in the lab, but also became a good friend who gave me encouragement and support throughout my PhD.

To my family back home – mom, dad and "little" sister. I know that moving so far away was not easy for you and I am sorry I missed so many important moments. Thank you for your unlimited support that always help me to keep going.

Last but not least, I would like to thank my fiancé Eduardo who helped me after we moved to Canada when I could barely speak English. I am thankful for your support in so many challenging moments in the past few years. Your strength and love were fundamental for my success in completing this PhD. Thank you for sharing your life with me and I am looking forward to our next adventures together!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANI	Average Nucleotide Identity
ANOVA	Analysis of variance
ATCC	American Type Culture Collection
BMI	Body mass index
BSM	Bifidus selective media
BV	Bacterial vaginosis
CDC	Centers for Disease Control and Prevention
COG	Cluster of Orthologous Group
CRISPR	Clustered regularly interspaced short palindromic repeats
CSB	Columbia agar 5% sheep blood
CST	Community State Type
DNA	Deoxyribonucleic acid
FDR	False discovery rate
GIT	Gastro-intestinal tract
HMO	Human milk oligosaccharides
LAB	Lactic acid bacteria
mB	Modified Brucella agar media
MIC	Minimum inhibitory concentration
NCBI	National Center for Biotechnology Information
NICU	Neonatal intensive care unit
OTU	Operational Taxonomic Units
PCoA	Principal coordinates analysis
PCR	Polymerase chain reaction
PTB	Preterm birth
qPCR	Quantitative PCR
rRNA	Ribosomal ribonucleic acid
SNP	Single nucleotide polymorphism
sPTB	Spontaneous preterm birth
TMB	Modified Brucella agar media containing tetramethylbenzidine
tRNA	Transfer ribonucleic acid
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
UT	Universal Target
WHO	World Health Organization

CHAPTER 1. Introduction and literature review

1.1. Introduction

The human body is estimated to contain as many bacterial cells as human cells, which represents a magnitude of 10^{13} bacteria (Sender et al., 2016). These bacteria live on external and internal surfaces including the gastrointestinal tract, vagina, breast, skin, and oral cavity. These microbial communities have the potential to impact human health by influencing processes related to development, nutrition, immunity, and resistance to pathogens (Bäckhed et al., 2004; Cho and Blaser, 2012; Maslowski and Mackay, 2011; Hooper and Gordon, 2001; Li et al., 2012).

In the lower genital tract, the complex bacterial community that resides in the human vagina is referred to as the vaginal microbiome, which has a significant role in women's reproductive health and neonatal health. Imbalances in this microbiota, also referred to as "dysbiosis", may lead to several negative conditions like bacterial vaginosis (BV), increased susceptibility to pelvic inflammatory disease (Ness et al., 2004; Peipert et al., 1997), and increased transmission of sexually transmitted pathogens like *Chlamydia trachomatis*, *Neisseria gonorrhoeae* and HIV (Wiesenfeld et al., 2003; Borgdorff et al., 2014; Buvé et al., 2014).

The current definition of a healthy vaginal microbiota in reproductive aged women is to be colonized by *Lactobacillus* spp.. Known mechanisms by which lactobacilli affect pathogen growth include adherence to mucus, developing a barrier that prevents pathogen colonization, and production of antimicrobial compounds such as lactic acid, hydrogen peroxide, bacteriocins, and biosurfactants (Boris and Barbés, 2000). Production of lactic acid is thought to be the hallmark characteristic of lactobacilli for maintaining vaginal microbiome homeostasis (O'Hanlon et al., 2013; Conti et al., 2009; Juárez Tomás et al., 2003; Aldunate et al., 2013; Boskey et al., 1999, 2001).

Bacterial vaginosis is the most common gynaecological condition in reproductive aged women worldwide and is characterized by a shift from healthy lactobacilli to a polymicrobial, anaerobic overgrowth of several species including *Gardnerella vaginalis*, *Bacteroides* spp., and *Mobiluncus* spp., (Sobel, 2000; Oakley et al., 2008; Hillier, 1993). The current gold standard method for diagnosis of BV is the Nugent score (Nugent et al., 1991), which is based on a scoring

system that reflects the relative abundance of bacterial cell morphotypes in Gram-stained vaginal smears. Although quick and inexpensive, the Nugent score does not provide reliable identification and quantification of bacteria at the species level.

In pregnancy, microbiome dysbiosis and BV have been associated with several negative reproductive outcomes of which preterm birth is the most common (Leitich et al., 2003; Meis et al., 1995). Nevertheless, no single bacterium or group of bacteria have yet been causally linked to prematurity, which emphasizes the complex multifactorial nature of this condition. In addition, understanding of the causes of microbiome imbalances in either pregnant or non-pregnant women is limited and the mechanisms underlying changes in the composition of the vaginal microbiota remain poorly characterized, which highlights the need for better methods to investigate the vaginal microbial community in health and “disease”.

The development of culture-independent methods has facilitated comprehension of the composition of vaginal microbial communities. As a result of application of DNA sequence based methods, there is a growing appreciation of "atypical" (i.e. non-*Lactobacillus* dominated) microbiomes in healthy women, such as those dominated by *Bifidobacterium* spp. (Hyman et al., 2005; Albert et al., 2015; Shipitsyna et al., 2013; Ravel et al., 2013). Based on the simplistic definition of healthy vaginal microbiome (*Lactobacillus*-dominated), those *Bifidobacterium*-dominated microbiomes are often grouped with microbial profiles that are recognisably unhealthy/disturbed, which might misrepresent their importance in also promoting a healthy vaginal microbiome.

Bifidobacteria have been extensively studied due to their health promoting characteristics in the gastro-intestinal tract, such as immune modulation (Fanning et al., 2012; Hart et al., 2004), production of bacteriocins (Martinez et al., 2013), improved gut metabolism (Sugahara et al., 2015), and inhibition of pathogens (Bernet et al., 1993; Arboleya et al., 2011; Fukuda et al., 2011). In newborns, bifidobacteria play an important role as one of the primary colonizers of the gastrointestinal tract (Arrieta et al., 2014) and have critical implications for early healthy immune and metabolic maturation (Rautava et al., 2012). Although bifidobacteria play an important role in gut health, their contribution to promoting a healthy vaginal microbiome has not yet been investigated. It is also unclear if vaginal bifidobacteria are an adapted subpopulation or if they are ubiquitous bacteria that can colonize both the gut and the vagina.

Moreover, bifidobacteria are likely vertically transmitted from mother to infant as several studies have demonstrated the influence of maternal microbiotas (vagina, gut, and breast milk) on the neonatal gut microbiome (Matamoros et al., 2013; Bäckhed et al., 2015; Makino et al., 2013; Dominguez-Bello et al., 2010; Gabriel et al., 2017; Gomez-Gallego et al., 2016; Hunt et al., 2011; Milani et al., 2015; Duranti et al., 2017; Pannaraj et al., 2017). The overlapping occurrence of bifidobacteria species, among other microorganisms, in different body sites is one of the challenges in studying vertical transmission, and the specific contribution of each microbial community in mother-to-infant bifidobacteria transmission remains in discussion. Research investigations of vaginal and gut bifidobacteria using high-resolution methods (e.g. genome sequence comparison) might help to elucidate some aspects of bifidobacteria ecology and indirectly illuminate aspects of mother-infant bifidobacteria transmission, an important event for neonatal gut microbiota development.

1.2. The vaginal microbiome

The healthy microbiota

The human vaginal microbiome is a complex bacterial community that has an important role in women's reproductive health. Imbalances in this microbiota may lead to increased susceptibility to sexually transmitted diseases, bacterial vaginosis (BV), and negative reproductive outcomes. Currently, a healthy vaginal microbiota in reproductive age women is defined as being colonized by *Lactobacillus* spp., with the most commonly detected species being *L. crispatus*, *L. gasseri*, *L. jensenii* and *L. iners* (Ravel et al., 2011). Lactobacilli can inhibit pathogens and contribute to microbiome homeostasis by forming a barrier on the vaginal epithelium through self-aggregation and adherence to mucus and epithelial cells. They can also compete with other bacteria for adherence and block their attachment to epithelial cells (Boris *et al.*, 1998; Osset *et al.*, 2001; Phukan *et al.*, 2013; Atassi *et al.*, 2006).

Lactobacilli also produce several metabolites that interfere with pathogens, such as lactic acid, hydrogen peroxide, bacteriocins and biosurfactants. Lactic acid is present in high concentrations (~56-111 mM) in the healthy vagina and is considered a hallmark characteristic of the protective lactobacilli due its microbicidal and virucidal effects (O'Hanlon et al., 2013; Conti et al., 2009; Juárez Tomás et al., 2003; Aldunate et al., 2013). Lactic acid is mainly responsible for

lowering the vaginal pH to ~3.5-4.5 and this acidic condition has been shown to reduce viability of several urogenital pathogens (Boskey et al., 1999, 2001; O'Hanlon et al., 2013). Other microbicidal mechanisms of lactic acid include disruption of the outer membrane of Gram-negative bacteria (Alakomi et al., 2000) and inhibition of viral replication and intracellular antigen synthesis (Conti et al., 2009).

Bacteria, including lactobacilli, are the primary source of lactic acid in the vagina, with some species being able to produce both lactic acid enantiomers (D and L), as opposed to human cells that only make the L form (Boskey et al., 2001). Although both forms of lactic acid have been shown to be microbicidal at physiological concentrations, L-lactic acid has been reported as having greater microbicidal and virucidal activity than D-lactic acid, which suggests that the effect of lactic acid involves more than acidity alone (Fayol-Messaoudi et al., 2005; Aldunate et al., 2013; Leitch and Stewart, 2002).

For decades hydrogen peroxide has been claimed to inhibit growth of opportunistic bacteria that do not have defensive mechanisms like production of catalase or peroxidases (Eschenbach et al., 1989; Strus et al., 2006; Klebanoff et al., 1991). Recent studies, however, have suggested that the level of hydrogen peroxide in vivo is much lower than the necessary concentration to be microbicidal (O'Hanlon et al., 2011). It has also been demonstrated that hydrogen peroxide at high concentrations is equally toxic to pathogens and lactobacilli, and that cervicovaginal fluid and semen inactivate the microbicidal activity of hydrogen peroxide produced by lactobacilli (O'Hanlon et al., 2010). Taken together, these recent findings challenge the claimed protective role of hydrogen peroxide and should be taken into account to better define the critical characteristics of health-promoting lactobacilli.

Bacteriocins and bacteriocin-like substances are antimicrobial peptides produced by certain bacteria to inhibit the growth of other bacteria. Bacteriocins secreted by vaginal lactobacilli have been reported to be active against several genital pathogens (Sabia et al., 2014; Stoyancheva et al., 2014; Ruíz et al., 2012). Lactobacilli also produce biosurfactants: substances composed of proteins and polysaccharides that reduce surface tension. In the vaginal environment, biosurfactants are thought to interfere with bacteria adhesion and subsequent biofilm formation (Satpute et al., 2016; Walencka et al., 2008).

Shifts in vaginal microbial composition occur naturally throughout a woman's life, mostly driven by hormonal changes. The sex steroid hormone estrogen stimulates the deposition of glycogen and the thickening of the vaginal epithelium, a stratified, squamous non-keratinized layer (Cruickshank and Sharman, 1934; BMJ, 1943). Glycogen is the main carbohydrate utilized by *Lactobacillus* for the production of lactic acid, which contributes to the protective effect of a low vaginal pH (Mirmonsef et al., 2014; Cruickshank, 1934; Boskey et al., 1999). At birth, newborn females have a stratified squamous vaginal epithelium resembling the adult epithelium due to maternal estrogen exposure. In childhood, the vaginal epithelium is thin and pH is near neutral due to the lack of glycogen. In this period, the vaginal microbiota is comprised of microorganisms similar to those found in the skin and gut microbiota like *Staphylococcus* and *Enterococcus* (Csángó, 1982; Gregoire et al., 1971; Cruickshank and Sharman, 1934). As estrogen levels rise during puberty, increasing amounts of glycogen are deposited in the vaginal epithelium. This acidic, glycogen rich, lactobacilli-dominated environment theoretically persists throughout a woman's reproductive years. Since estrogen is directly related to glycogen levels in the vaginal epithelium, it has been hypothesized that the vaginal microbiota changes through a woman's menstrual cycle in response to fluctuations in estrogen levels. Culture-based studies have led to variable conclusions, with some studies reporting no changes in the microbiota throughout the menstrual cycle (Wilks and Tabaqchali, 1987), and others describing a greater proportion of non-*Lactobacillus* species during menses (Eschenbach et al., 2000). Recent culture-independent studies have demonstrated that although shifts in the vaginal microbiota are relatively common, those changes are unlikely to be associated with menstrual phases (Chaban et al., 2014; Hickey et al., 2013). After menopause, the decrease in estrogen is accompanied by a decline in glycogen in the vaginal epithelium, which results in changes in the epithelial structure and a decline in lactobacilli (Cruickshank and Sharman, 1934).

The vaginal microbiome in dysbiosis

Shifts in the vaginal microbial composition are expected to occur due to aging, but changes in the vaginal microbiota during the reproductive years are also common. Understanding the causes of such variations, however, is limited and the mechanisms underlying changes in the composition of the vaginal microbiota remain poorly characterized. It is known that imbalances of

the vaginal microbiome, or dysbiosis, may lead to negative conditions including bacterial vaginosis (BV), which is the most common gynaecological condition in reproductive aged women worldwide (Sobel, 1997; Koumans et al., 2007; Kenyon et al., 2013). BV is currently characterized by a relatively low abundance of *Lactobacillus* spp. accompanied by polymicrobial anaerobic overgrowth, including species such as *Gardnerella vaginalis*, *Prevotella* spp., *Bacteroides* spp., *Mobiluncus* spp., and *Mycoplasma hominis* (Sobel, 2000; Oakley et al., 2008; Hillier, 1993). Typical signs of inflammation such as redness, swelling and pain are absent in BV, but a pro-inflammatory response might occur as a result of BV-associated bacteria colonization (Mitchell and Marrazzo, 2014).

Bacterial vaginosis has been historically reported as 'nonspecific vaginitis' to describe cases in which vaginitis-like symptoms were not attributable to the known etiologic agents *Trichomonas vaginalis* or *Candida albicans* (Hill, 1993). In 1955, Gardner and Dukes proposed that a new isolated rod named *Haemophilus vaginalis* (currently known as *Gardnerella vaginalis*) was the aetiologic agent of 'nonspecific vaginitis', which was then replaced by the term 'Haemophilus vaginalis vaginitis' (Gardner and Dukes, 1955). Later, the idea that *Gardnerella vaginalis* was the sole agent of nonspecific vaginitis was criticized and the condition was recognized as multifactorial since other anaerobic bacteria are often associated with microbiome dysbiosis, and *G. vaginalis* is also detected in asymptomatic women with "healthy" microbiotas (Spiegel et al., 1980; Amsel et al., 1983). In the 1980's, the term 'bacterial vaginosis' was gradually implemented to better reflect the bacterial diversity of the condition, and to avoid the misleading suggestion that BV causes inflammation (as opposed to vaginitis), which is absent during vaginal dysbiosis (Holmes et al., 1981; Easmon et al., 1992).

Although BV does not have a specific etiologic agent, *G. vaginalis* is still considered a hallmark bacterium involved in this condition. *G. vaginalis* was first isolated by Leopold in 1953 and described as a small, Gram-negative, non-motile, non-capsulated, pleomorphic rod-shaped bacterium. In 1955, Gardner & Dukes isolated the same organism from women with non-specific vaginitis (today known as BV) and named it '*Haemophilus vaginalis*'. Later, Zinnemann & Turner (1963) recommended the reclassification of *H. vaginalis* to the genus *Corynebacterium* based on morphological parameters, and they renamed the organism as *Corynebacterium vaginale*. In 1980, Greenwood & Pickett used DNA-DNA hybridization, biochemical and cell wall analysis and electron microscopy to show that *H. vaginalis*, or *C. vaginale*, did not belong to the genus

Haemophilus. They created the genus *Gardnerella* and renamed the organism after Hermann L. Gardner: *Gardnerella vaginalis*. In the same year, Piot and colleagues confirmed their study and supported the name *G. vaginalis*, which has been used since then (Piot et al., 1980).

Phenotypic characteristics typically used for bacterial identification vary significantly among *G. vaginalis* strains; overall, they present beta-hemolysis on human blood agar, negative catalase reaction, positive hippurate hydrolysis and starch hydrolysis (Piot et al., 1982). *G. vaginalis* also have virulence factors that are thought to contribute to microbiome imbalances, including adherence to epithelial cells, biofilm production (Patterson et al., 2010), cytolysin activity (Cauci et al., 1993), and sialidase activity (Hardy et al., 2017). A genotyping method based on amplified ribosomal DNA restriction analysis (ARDRA) has been proposed to better identify biotypes of *G. vaginalis* that might share similar characteristics (Ingianni et al., 1997), but inconsistent reports failed to correlate biotype groups with specific phenotypes (Paramel Jayaprakash et al., 2012; Lopes dos Santos Santiago et al., 2011; Pleckaityte et al., 2012). More recently, four distinct subgroups of *G. vaginalis* (A, B, C and D) have been proposed based on their cpn60 sequences. These subgroups are supported by whole genome sequence data and have been shown to correlate with sialidase activity, which might have implications in their role in BV (Paramel Jayaprakash et al., 2012; Schellenberg et al., 2016).

A mounting body of evidence has demonstrated that vaginal microbiome dysbiosis or BV is strongly associated with several negative conditions including increased susceptibility to pelvic inflammatory disease (Ness et al., 2004; Peipert et al., 1997), and increased transmission of several sexually transmitted pathogens including *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, *Trichomonas vaginalis*, human immunodeficiency virus (HIV) and herpes simplex virus type 2 (Allsworth and Peipert, 2011; Borgdorff et al., 2014; Brotman, 2011; Gallo et al., 2012; Martin et al., 1999; Taha et al., 1998; Wiesenfeld et al., 2003; Buvé et al., 2014; Bautista et al., 2016; Ziklo et al., 2016; Cherpes et al., 2003).

The causes of BV/dysbiosis are not fully understood, but host and environmental factors are thought to play important roles in influencing the composition of microbial communities. Smoking (Ryckman et al., 2009; Brotman et al., 2014a), intercourse, douching, some sexual practices (Cherpes et al., 2008) and use of feminine hygiene products (Fashemi et al., 2013) are often associated with a decrease in vaginal lactobacilli and a shift toward an imbalanced

microbiome. Moreover, reports showing varying incidence of microbiome dysbiosis among women of different ethnicities suggest host factors also play a role in shaping the microbiota. Overall, there is lower incidence of vaginal communities dominated by *Lactobacillus* in black women than in Caucasian women (Zhou et al., 2007; Ravel et al., 2011). Hispanic and black women are also more likely to have higher vaginal pH compared to Asian and white women, which reflects the higher prevalence of communities not dominated by *Lactobacillus* spp. (Ravel et al., 2011). These differences cannot be explained by external factors like personal hygiene and sexual behaviours alone (Goldenberg et al., 1996), which emphasizes the role of host factors in driving microbial community structure; differences in innate and adaptive immune systems, composition and quantity of vaginal secretions, and ligands on epithelial cell surfaces might be involved (Ravel et al., 2011).

The gold standard method for diagnosis of BV is the microbiological criteria introduced by Nugent, Krohn, & Hillier (1991). The Nugent score is based on a numerical scoring system (0–10), where a score of 0–3 is considered healthy, 4–6 as intermediate, and 7–10 as indicative of BV. The score reflects the relative abundances of three kinds of bacterial cell morphotypes in Gram-stained vaginal smears: large Gram-positive rods (*Lactobacillus*), small Gram-negative and Gram-variable rods (*Gardnerella vaginalis*/*Bacteroides* spp.), and curved Gram-negative rods (*Mobiluncus*) (Nugent et al., 1991). Bacterial vaginosis can also be diagnosed through the clinical criteria established by Amsel and colleagues in 1983, in which BV is characterized by the presence of at least three of the following four criteria: 1) thin, gray/white discharge; 2) malodorous "fishy" discharge upon adding 10% potassium hydroxide; 3) high vaginal pH (>4.5), and 4) identification of vaginal epithelial cells heavily coated with bacteria (i.e., "clue cells") (Amsel et al., 1983).

Current treatment strategies for BV include the administration of antibiotics either orally or topically. The Centers for Disease Control and Prevention (CDC) recommends treatment only for women with symptoms, which accounts for a low proportion of BV cases based on Nugent score (Koumans et al., 2007). Treatment regimens consist of metronidazole 500 mg orally twice a day for 7 days; or metronidazole gel 0.75%, one full applicator (5g) intravaginally, once a day for 5 days; or clindamycin cream 2%, one full applicator (5g) intravaginally at bedtime for 7 days (CDC - Centers for Disease Control and Prevention, 2015b). Cure of BV is commonly defined as failure to meet Amsel's criteria (≥3 symptoms present) after treatment (Verstraelen and Verhelst, 2009). The main disadvantage of antimicrobial therapy is that it usually only provides a short term

relief from BV symptoms, typically assessed 7-14 days after starting treatment (Sobel, 2009). In the long term, BV cure rates vary widely, ranging from 60-90% after one month of treatment (Larsson and Forsum, 2005; Koumans et al., 2002) to 42% within a year after treatment (Bradshaw et al., 2006). Although BV recurrence is relatively common, most cases likely represent a relapse rather than a reinfection, especially if symptoms reappear shortly after treatment (Sobel, 2009).

One of the main contributors to BV therapy failure is bacterial antibiotic resistance. Many non–spore-forming, Gram-positive anaerobic rods are resistant to metronidazole and clindamycin; this group of bacteria includes *Propionibacterium* and *Actinomyces* (*Atopobium, Mobiluncus*) (Hall and Copsey, 2015). *Prevotella* (Alauzet et al., 2010) and *Gardnerella vaginalis* have also been reported to be resistant to metronidazole (Aroutcheva et al., 2001; Kharsany et al., 1993). Vaginal biofilms also contribute to antibiotic treatment failure since they allow bacteria to reach much higher concentrations than in luminal fluids and prevent/delay antimicrobial agents to reach the bacteria (Swidsinski et al., 2008; Donlan and Costerton, 2002). *Gardnerella vaginalis* and *Atopobium vaginae* are often together in vaginal biofilms (Swidsinski et al., 2005) and it has been reported they are only temporarily suppressed by antibiotics and rapidly regain activity after antimicrobial treatment (Swidsinski et al., 2008).

Pregnancy and the vaginal microbiome

In past decades, several culture based and culture-independent studies have described the vaginal microbiome of healthy, reproductive aged, non-pregnant women, but studies of the microbiome in pregnancy are still scarce. Pregnancy is associated with a variety of physiological events that have the potential to influence the structure of the vaginal microbial community, including increased sex steroid hormone levels (Tulchinsky et al., 1972), host immune response modulation (Jamieson et al., 2006; Mor and Cardenas, 2010), altered immune-physicochemical properties of the cervical mucus (House et al., 2009; Luo et al., 2000; Lee et al., 2011a), as well as behavioural changes such as reduced drinking and smoking (Crozier et al., 2009). Nevertheless, the precise mechanisms by which these factors influence the vaginal microbiome are not fully understood.

Considering the hormonal and physiological changes associated with pregnancy, a few culture-independent studies have compared the vaginal microbiome of healthy pregnant and non-

pregnant women. The microbiomes of pregnant women are overall less diverse and less rich when compared to those of non-pregnant women; they also have higher abundance of *Lactobacillus* spp. and are more stable (Aagaard et al., 2012; Walther-António et al., 2014; Romero et al., 2014c; MacIntyre et al., 2015). After delivery, post-partum microbiomes become less *Lactobacillus* dominant and more rich and diverse (i.e. more similar to the microbiomes of non-pregnant women), which supports the idea that pregnancy has a transient effect on the vaginal microbial community (MacIntyre et al., 2015). These differences are likely driven by the increased levels of estrogen during pregnancy that lead to increased glycogen deposition when compared to non-pregnant women (Cruickshank and Sharman, 1934; BMJ, 1943). As mentioned previously, glycogen is utilized by *Lactobacillus* to produce lactic acid and contributes to the protective role of low vaginal pH (Mirmonsef et al., 2014; O'Hanlon et al., 2013).

Notably, most studies of the vaginal microbiome in pregnancy have only included study populations of American (United States) or European origin, and descriptions of the vaginal microbiome of pregnant women from other regions, including Canada, are lacking. Moreover, these studies are based on the sequencing of the 16S rRNA gene, a barcode that does not provide enough resolution to resolve closely related organisms, particularly some members of the *Bifidobacteriaceae* family like *Bifidobacterium* and *Gardnerella* (Paramel Jayaprakash et al., 2012; Vermette et al., 2010; Verbeke et al., 2011; Lan et al., 2016; Hill et al., 2010).

One of the main motivations for studying the vaginal microbiome in pregnancy relates to previous reports of an association between BV/dysbiosis and negative reproductive outcomes including increased risk of post-abortion infection (Larsson et al., 1992, 2000; Hamark and Forssman, 1991), early (Ralph et al., 2015; Donders et al., 2000) and late miscarriage (Llahi-Camp et al., 1996; Hay et al., 1994), chorioamnionitis (Gibbs, 1993; Hillier et al., 1988), postpartum endometritis (Watts et al., 1990; Jacobsson et al., 2002), preterm premature rupture of membranes (PPROM) (Parry and Strauss, 1998; Hillier et al., 1995) and preterm birth (Leitich et al., 2003; Meis et al., 1995).

Approaches to studying the vaginal microbiome

The structure of vaginal microbial communities has been investigated over past decades using a wide range of culture-dependent and -independent techniques. Culture based studies rely

on the ability of microorganisms to grow in laboratory culture media, and therefore, they are likely to underestimate the diversity of the vaginal microbiota since many fastidious bacteria will not grow in the culture media. Culture-based studies, however, were essential in the early days when vaginal dysbiosis was a completely uncharacterized condition. By using culture methods, Gardner and Dukes were able to isolate a bacterium from women with 'nonspecific vaginitis' which led to the proposition that *G. vaginalis* was the etiologic agent of BV (Gardner and Dukes, 1955). It is known today that other anaerobic bacteria also play a role in BV, but those initial reports were an important step towards a better understanding of the vaginal microbiota. Subsequent culture-based studies not only contributed to a better understanding of microbiota diversity, but also improved knowledge regarding the role of pH, lactic acid, hydrogen peroxide and lactobacilli in the microbiome (Wilks and Tabaqchali, 1987; Spiegel et al., 1980; Eschenbach et al., 1989; Csángó, 1982; Goldenberg et al., 1996; Onderdonk et al., 1986). Culturing microorganisms also provides the opportunity for establishing a culture collection that can be used to investigate virulence factors or health promoting effects through in vitro assays or whole genome sequencing.

In 1983, Amsel and colleagues proposed the clinical criteria (vaginal pH>4.5; vaginal discharge; amine odour upon KOH addition; clue cells in Gram stain) as the first attempt to formally diagnose BV. BV symptoms might be subtle and subjective to interpret; also, *G. vaginalis* and other anaerobic bacteria are often isolated from the vaginal microbiota of asymptomatic women. Thus, in 1991, Nugent, Krohn & Hillier proposed microbiological criteria (Nugent score) to diagnose BV based on Gram-stained vaginal smears. The Nugent score became the gold standard method and it has been used since then in studies investigating the vaginal microbiota. The Nugent score is inexpensive, quick and widely available in many research and clinic facilities, but it does not provide reliable quantification and identification of bacteria at the species level.

The development of culture-independent methods has greatly facilitated the investigation of the vaginal microbiota, and improved the resolution of vaginal microbiome characterization in a variety of cohorts. Some of the fingerprint techniques that have been used to analyse the composition of the vaginal microbiota include PCR-denaturing gradient gel electrophoresis (DGGE) (Ling et al., 2010; Gao et al., 2013; Burton et al., 2003a) and terminal restriction fragment length polymorphism (tRFLP) (Verstraelen et al., 2009; Thies et al., 2007; Zhou et al., 2007). These approaches are relatively inexpensive and give an overview of the species present in a sample, but do not provide a high resolution analysis (Walker, 2016). Other molecular techniques

like end-point PCR (DiGiulio et al., 2010; Fredricks et al., 2007) and quantitative PCR (Menard et al., 2010; Fredricks et al., 2009; Jespers et al., 2012) are also widely used due to their specificity to detect organisms at the genus or species level, but they do not give information about new species since only known targets can be detected.

Advances in DNA sequence-based methods have revolutionized the study of the human microbiome. One common methodology is to carry out surveys of universal marker genes, such as 16S rRNA or cpn60, which provide a broad representation of the bacterial species present in a sample. The universal target genes are PCR-amplified generating a mixed pool of PCR amplicons that can be used to construct clone libraries for subsequent sequencing, or alternatively, be directly sequenced en masse (Walker, 2016). Clone libraries containing 16S rRNA amplicons (Oakley et al., 2008; Hyman et al., 2005) or cpn60 amplicons have both been used as an approach to study the vaginal microbiota (Hill et al., 2005b; Schellenberg et al., 2011).

Direct amplicon sequencing is currently the most common approach to study the composition of the vaginal microbiota, either based on the PCR amplification of 16S rRNA gene (Hummelen et al., 2010; Ravel et al., 2013; MacIntyre et al., 2015; Hickey et al., 2013; Aagaard et al., 2012; Shipitsyna et al., 2013; Huang et al., 2015; Gajer et al., 2012; Ravel et al., 2011) or the cpn60 gene (Schellenberg et al., 2011, 2009; Chaban et al., 2014; Albert et al., 2015). As a result of application of DNA sequence based methods to characterize the composition of the vaginal microbiome, there is a growing appreciation of "atypical" (i.e. non-*Lactobacillus* dominated) microbiomes in healthy women, such as those dominated by *Bifidobacterium* spp. that have been identified in several cohorts of healthy women (Hyman et al., 2005; Albert et al., 2015; Shipitsyna et al., 2013; Ravel et al., 2013). Atypical healthy microbiomes are often grouped with microbial profiles that are recognisably unhealthy/disturbed, and they are also often overlooked and generally not further investigated in vaginal microbiome studies.

Recently developed "–omics" approaches have great potential to provide in-depth analysis of any microbial community, and several studies have used metagenomics/whole-genome shotgun sequencing (Prince et al., 2016; Zevin et al., 2016), transcriptomics (Macklaim et al., 2013), proteomics (Borgdorff et al., 2015; Klatt et al., 2017; Zevin et al., 2016), and metabolomics (McMillan et al., 2015; Srinivasan et al., 2015) to investigate the vaginal microbiome. Although

promising, sequencing costs and scarce bioinformatics tools for "big data" analysis are still limiting factors for the –omics approaches (Martin and Marrazzo, 2016).

1.3. Preterm birth

Definition

Preterm birth (PTB) is defined by WHO as delivery before 37 weeks of gestational age (WHO *et al.*, 2012). PTB can be further sub-categorized into extremely preterm ($\leq 27^{+6}$ weeks^{+days}), very preterm (28 to 31^{+6}), and late preterm (32 to 36^{+6}) (Quinn et al., 2016). Eleven percent of all live births worldwide are preterm, which is estimated to be the cause of 35% of the world's neonatal deaths and represents 3.1 million deaths per year (Blencowe et al., 2012). In Canada, the average preterm birth rate among live births was 7.8% in 2013, with Nunavut (12.6%) and Prince Edward Island (7.1%) having the highest and lowest preterm rates, respectively (Statistics Canada, 2016). Preterm birth complications are responsible for an increased risk of cardiovascular disorders, respiratory distress syndrome, neurodevelopmental disabilities and learning difficulties in children who are born prematurely in comparison with children born at term (Behrman and Butler, 2007).

Hypothesis of mechanisms

Pathogenesis of preterm labour is not yet fully understood, but current hypotheses involve early idiopathic activation of the normal labour process or pathological insults (Goldenberg et al., 2008). Both term and preterm labour involve similar events as part of the "common pathway" of labour, which includes increased uterine contractility, cervical dilatation, rupture of the chorioamniotic membranes, and changes in the concentrations of hormones corticotrophin-releasing factor and cortisol (Romero et al., 2006). Two pathways are thought to regulate the timing of parturition: an endocrine cascade comprising the fetal hypothalamic–pituitary–adrenal–placental axis, and a mechanical pathway in which fetal growth enhances uterine wall tension, inducing biochemical and molecular changes in the myometrium (Shynlova et al., 2013). These changes in the myometrium are thought to shift the signalling from anti-inflammatory to pro-inflammatory pathways, which involves several events like progesterone withdrawal (Renthal et

al., 2013), cervical ripening, and decidual activation (Romero et al., 2014a) that in turn lead to the 'common pathway' of labour. Thus, a coordinated inflammatory response, including infiltration of neutrophils and macrophages in the myometrium (Thomson et al., 1999), has been implicated to have a central role in term labour (Bollopragada et al., 2009; Dudley et al., 1996).

Known pathological mechanisms implicated in preterm labour include: 1) Failure of physiologic transformation (disappearance of the normal muscular and elastic structures of the arteries and their replacement by fibrinoid material) that can cause maternal lesions and uteroplacental ischaemia predisposing to preeclampsia (Kim et al., 2003); 2) Uterine overdistension (as a result of multiple gestations), polyhydramnios (an excessive amount of amniotic fluid) or macrosomia (large fetal weight) that might induce expression of gap junction proteins, IL-8 and collagenase production facilitating cervical ripening and myometrium activation (Behrman and Butler, 2007); 3) Cervical insufficiency, a syndrome that leads to cervical ripening and is caused by congenital disorders (i.e. hypoplastic cervix), surgical trauma (i.e. conization resulting in substantial loss of connective tissue), or traumatic damage to the structural integrity of the cervix (i.e. repeated cervical dilatation associated with termination of pregnancy) (Romero et al., 2006); 4) Progesterone withdrawal mediated by local changes in metabolism or alterations in progesterone receptors isoforms/coactivators (Romero et al., 2014a); 5) Disruption of maternal-fetal tolerance caused by abnormalities in the recognition and adaptation to fetal antigens (Lee et al., 2011b).

Preterm birth is considered a complex multi-factorial condition and several risk factors have been associated with early parturition. Such factors include history of prematurity (Mercer et al., 1999), black ethnicity (Kistka et al., 2007), close temporal proximity to a previous delivery (Conde-Agudelo et al., 2006), multiple gestations (Kurdi et al., 2004; Stock and Norman, 2010), short cervix (Iams et al., 1996), low and high maternal ages (Fraser et al., 1995; Jacobsson et al., 2004; Schempf et al., 2007), low maternal BMI (Han et al., 2011), tobacco use, heavy alcohol intake, illicit drug use (Behrman and Butler, 2007), and maternal medical disorders (e.g. thyroid disease, asthma, diabetes, hypertension) (Behrman and Butler, 2007). Maternal periodontal disease has also been implicated in prematurity, though the exact mechanism has not been elucidated. It is thought that periodontal infections can stimulate a local and systemic host response that activates the pro-inflammatory cascade involved in preterm birth (Bobetsis et al., 2006; Xiong et al., 2006).

Although studied extensively, some preterm cases remain unexplained for women with no pathological conditions or known risk factors.

An intrauterine ascending infection hypothesis has been proposed as an important contributor for preterm birth, and is supported by the fact that many organisms isolated from the amniotic fluid/membranes of women who had preterm delivery are also found in the lower genital tract of pregnant women (Gardella et al., 2004; Krohn et al., 1995; Hillier et al., 1988; Romero et al., 1989). By ascending from the vagina, bacteria can reach the uterus and decidua and stimulate a local inflammatory reaction through the production of proinflammatory cytokines and inflammatory mediators. They can also cross the chorioamnion into the amniotic cavity and stimulate the production of inflammatory mediators by local macrophages and other host cells. Eventually, the organisms might gain access to the fetus and initiate a systemic inflammatory response (Romero et al., 2006).

A large number of studies support the ascending infection hypothesis based on the strong association between intra-amniotic bacterial infection and preterm birth (Gardella et al., 2004; Hillier et al., 1988; Romero et al., 1989; Watts et al., 1992; Yoon et al., 2001; DiGiulio et al., 2008; Hillier et al., 1993; Jacobsson et al., 2005). Most intra-amniotic infections, however, are considered subclinical (Romero et al., 1989) since they do not meet criteria for clinical chorioamnionitis (fever, uterine fundal tenderness, maternal tachycardia >100/min, fetal tachycardia >160/min, and purulent or foul amniotic fluid) (Tita and Andrews, 2010), which might lead to an underestimation of the importance of infection as a contributor for preterm birth.

Preterm labour caused by an ascending infection is mediated by an inflammatory process. Microorganisms are recognized by toll-like receptors (TLRs) and other pattern recognition receptors that initiate a proinflammatory cascade mostly mediated by the transcription factor NF-kB (Lee et al., 2003). This cascade stimulates the production of cytokines (IL-1β, IL-6 and TNF-α), chemokines (IL-8) (Romero et al., 2006), prostaglandins (Mitchell et al., 1991, 1993), and proteases (Phillippe et al., 2001) that lead to the activation of the 'common pathway' of parturition (Romero et al., 2006; Agrawal and Hirsch, 2012). However, the precise timing between infection and initiation of labour might be difficult to estimate as the inflammatory cascade might be activated long before the signs of labour are apparent (Agrawal and Hirsch, 2012).

The presence of bacteria alone may not be sufficient to cause preterm labour as reports have shown that bacteria may be present in fetal membranes without necessarily causing an inflammatory response (Steel et al., 2005). This also indicates that not all dysbiotic microbiomes (simplistically defined as lack of lactobacilli) are involved in preterm birth since different bacteria have different virulence factors and might stimulate distinct inflammatory responses. Studies based on culture-dependent techniques have failed to provide in-depth understanding of the microbial community that is associated with preterm birth, and studies using high throughput sequencing technologies to address this question are still scarce.

Microorganisms and preterm birth

Gardnerella vaginalis

Several pathogens and BV-associated bacteria have been previously associated with early onset of labour. *G. vaginalis*, a hallmark bacteria in BV, is often detected in women who experienced preterm birth (Bretelle et al., 2015). *G. vaginalis* is often described as Gram-variable due to a very thin cell wall, but ultrastructural characteristics and chemical composition indicate they are in fact Gram-positive (Sadhu et al., 1989). *G. vaginalis* produce enzymes including sialidases that can degrade cervicovaginal mucus (Lewis et al., 2013) and therefore facilitate bacteria ascension to the uterus. Elevated sialidase levels have been strongly associated with early spontaneous preterm birth (Cauci and Culhane, 2011). Sialidases (neuraminidases) cleave the terminal sialic acid residues linked to oligosaccharide chains (Lewis and Lewis, 2012); by breaking down the sialic acid-containing mucus components in the vagina, *G. vaginalis* contributes to the depletion of protective host mucus barriers (Lewis et al., 2013). Other immune molecules present in the vaginal mucosal surface, including immunoglobulins (IgA, IgG), lysozyme, cytokines and cell surface receptors that are important to the ability of the reproductive mucosa to neutralize and eliminate pathogens, are also metabolized by sialidases (Lewis et al., 2012; Lewis and Lewis, 2012).

BV-associated bacteria

Several BV-associated bacteria have been associated with preterm birth. The most commonly detected bacteria associated with BV are *Megasphaera*, *Prevotella*,

Leptotrichia/Sneathia, Atopobium vaginae, Eggerthella-like bacterium, and bacteria in the order *Clostridiales* (BVAB1, BVAB2, BVAB3) (Fredricks et al., 2007). High levels of *Atopobium vaginae* have been associated with preterm delivery (Menard et al., 2010). *A. vaginae* is a fastidious, Gram-positive, anaerobic, cocci-shaped bacterium that often occurs together with *G. vaginalis* (Trama et al., 2008). Both bacteria trigger similar inflammation and vaginal epithelial innate immune responses mediated by activation of the proinflammatory transcription factor NF-κB in vaginal epithelial cells (Libby et al., 2008).

Prevotella are anaerobic, Gram-negative, pleomorphic, non-motile rods previously classified as *Bacteroides* that have also been associated with preterm birth (Holst et al., 1994; Callahan et al., 2017). *Prevotella* are often detected with *G. vaginalis* and a synergistic relationship between them has been described; ammonia produced by *Prevotella* enhances *G. vaginalis* growth, which in turn produces amino acids that are utilized by *Prevotella* (Pybus and Onderdonk, 1998). Similarly to *G. vaginalis*, *P. bivia* and *P. disiens* also exhibit sialidase activity (Briselden et al., 1992), and *P. bivia* is thought to trigger proinflammatory responses in the vaginal mucosa by stimulating the production of IL-8 and NF-κB activation (Fichorova et al., 2011), which might be involved in the role of *Prevotella* in prematurity. Some recent studies, however, have reported no association between specific taxa of BV-associated bacteria and preterm birth (Romero et al., 2014b; Stout et al., 2017). Interestingly, several BV-associated bacteria (*Gardnerella*, *Prevotella*, *Atopobium, Mobiluncus*) were associated with preterm birth in a low risk cohort that mostly included Caucasian and Asian participants; while only *Prevotella* was associated with prematurity in a high risk for preterm cohort mainly composed of African American women with history of preterm birth, which suggest that preterm–microbiota associations are population-dependent (Callahan et al., 2017).

Mollicutes

Mollicutes (from the Latin 'soft skin') is a class of organisms that includes the genera *Mycoplasma* and *Ureaplasma*. They are typically very small bacteria and have a trilayered external membrane, rather than a rigid cell wall, rendering them resistant to β-lactam antibiotics (Taylor-Robinson, 2007). The most common species detected in the human genital tract are *Mycoplasma hominis*, *Mycoplasma genitalium*, *Ureaplasma urealyticum* and *Ureaplasma parvum*, and they are often referred as 'genital mycoplasmas' (Taylor-Robinson and Lamont, 2011). Mollicutes have

been associated with preterm birth likely mediated by an increased proinflammatory response, but inconsistent findings from different studies still make it unclear if they are involved with early parturition. An association between preterm birth and the following species have been reported: *M. hominis* (Lamont et al., 1987; Donders et al., 2000; Kwak et al., 2014), *M. genitalium* (Hitti et al., 2010; Lis et al., 2015; Edwards et al., 2006), *U. urealyticum* (Gerber et al., 2003; Yoon et al., 1998) and *U. parvum* (Kataoka et al., 2006). Other studies, however, have suggested that preterm birth occurs independently of mycoplasma colonization (Lu et al., 2001; Oakeshott et al., 2004; Carey et al., 1992).

Some genital mycoplasmas might be considered part of the "normal" microbiota of some women. *M. hominis* is detected in the cervix or vagina of 20–50% of sexually mature, asymptomatic woman, ureaplasmas in 40–80% and *M. genitalium* in 0–5% (Taylor-Robinson and Jensen, 2010). While some of these organisms are often found in the lower genital tract of women, they are less prevalent (<5%) in the upper genital tract and chorioamnion of pregnant women in the absence of labour or membrane rupture (Tita and Andrews, 2010). Moreover, an association between BV and Mollicutes has been reported (Taylor-Robinson, 2017), which indicates that the role of Mollicutes in preterm birth might involve the presence of other bacteria or conditions (Capoccia et al., 2013; Cox et al., 2016). Clearly, further studies investigating the vaginal microbial community in pregnancy are needed to clarify the role of Mollicutes in preterm birth.

Chlamydia trachomatis

Chlamydia trachomatis is an obligate intracellular pathogen that causes ocular and genital infections (Brunham and Rey-Ladino, 2005). Chlamydial infections are the most common sexually transmitted infections, are often asymptomatic, and if untreated, can cause pelvic inflammatory disease and infertility in women, and urethritis in men (PHAC - Public Health Agency of Canada, 2014). Women with chlamydial infections are also more likely to experience preterm birth (Andrews et al., 2000; Claman et al., 1995; Liu et al., 2013; Mann et al., 2010; Schoeman et al., 2011) probably mediated by an inflammatory response (Rours et al., 2011), although some studies have found no association between chlamydia and preterm birth (Andrews et al., 2006b; Hitti et al., 2010; Edwards et al., 2006). These inconsistent reports indicate that infection by chlamydia might be influenced by other factors such as the vaginal microbiota, that would affect the ability of chlamydia to cause, or not, preterm birth.

During chlamydia infection, the host immune response stimulates the production of the cytokine INF-γ. One of the roles of INF-γ is to upregulate the host enzyme IDO (indoleamine 2,3-dioxygenase), which converts tryptophan to kynurenine. Tryptophan is an essential amino acid for *C. trachomatis* and under this stressful condition of tryptophan depletion, the *Chlamydia* developmental cycle is disrupted driving the bacteria into its noninfectious form, known as reticulate bodies (RB) (AbdelRahman and Belland, 2005; Ziklo et al., 2016). To overcome the tryptophan depletion, it has been hypothesized that *C. trachomatis* can use its *trpBA* genes to rescue tryptophan from indole, which, in turn, is a product of anaerobic members of the vaginal microbiota, but not produced by lactobacilli (Ziklo et al., 2016). In vivo examination of women infected with *C. trachomatis* and BV revealed high levels of indole in their vaginal secretions (Lewis et al., 2014). This demonstrates how a disturbed microbiome can increase susceptibility to chlamydial infections, which in turn might influence whether chlamydia might cause preterm birth or not.

Neisseria gonorrhoeae

Neisseria gonorrhoeae is the second most commonly reported sexually transmitted pathogen in Canada (PHAC - Public Health Agency of Canada, 2014). It causes cervicitis, pelvic inflammatory diseases and infertility in women, and urethritis and infertility in men (CDC - Centers for Disease Control and Prevention, 2015a). Gonorrhea is currently treated with antimicrobial therapy, however there is a growing concern worldwide about widespread antimicrobial resistance (Wi et al., 2017). Gonorrhea has also been associated with preterm birth, but like mycoplasmas and chlamydia, the role of *N. gonorrhoeae* in preterm birth is still unclear based on inconsistent reports (Johnson et al., 2011; Liu et al., 2013; Mann et al., 2010; Heumann et al., 2017; Edwards et al., 2006; Hitti et al., 2010). It is known that lactic acid acidification or adherence inhibition mediated by lactobacilli strongly inhibits *N. gonorrhoeae*, which indicates that BV or lack of lactic acid–producing bacteria in the vaginal microbiota increases susceptibility to gonorrhea infection (Graver and Wade, 2011; Breshears et al., 2015; Spurbeck and Arvidson, 2008). Like chlamydia, it is possible that gonorrhea infection is influenced by the presence of specific organisms from a disturbed microbiota, which would affect *N. gonorrhoeae* ability to cause infection and preterm birth.

1.4. The genus *Bifidobacterium*

History

Bifidobacteria were first described by Henry Tissier (1889, 1900), who isolated a bacterium from breast-fed infant feces and named it *Bacillus bifidus communis*, or simply, *B. bifidus*. They were described as rods with ends split to give the characteristic Y- shape, which led to the designation of "bifid". *Bifidobacterium* have a complex taxonomic history; *B. bifidus* was initially assigned to the genus *Bacillus* (and later to the genus *Lactobacillus*) due morphological and biochemical similarities to *Bacillus acidophilus* (Poupard et al., 1973). In 1924, Orla-Jensen proposed the genus *Bifidobacterium* as a separate taxon for these organisms (Orla-Jensen, 1924). This designation was not generally accepted though, and the term *Bacillus*/*Lactobacillus* continued to be used to describe species that today are known to belong to *Bifidobacterium*. Numerous other names and re-classifications were also proposed in the following years. All the organisms' names proposed for *B. bifidum* are listed here in chronological order: *Bacillus bifidus communis, Bacillus bifidus, Bacteroides bifidus, Lactobacillus bifidus, Bifidobacterium bifidum, Bacterium bifidum, Tissieria bifida, Actinomyces befidus, Actinobacterium bifidum, Corynebacterium bifidum*, and *Cohnistreptothrix bifidus* (Poupard et al., 1973).

In 1957, Dehnart realized the existence of multiple biotypes of *Bifidobacterium* that could be differentiated based on their carbohydrate fermentation patterns. This drove the "discovery" of several *Bifidobacterium* species in the following years. In 1971, Reuter designated seven type strain species of *Bifidobacterium*, in addition to the known *B. bifidum: B. adolescentis, B. breve, B. infantis, B. lactentis, B. liberorum, B. longum*, and *B. parvulorum*. In 1974, Scardovi and Crociani proposed three type strains: *B. catenulatum*, *B. dentium, and B. angulatum.* And in 1979, Scardovi and colleagues proposed another four type strains: *B. cuniculi, B. choerinum, B. boum,* and *B. pseudocatenulatum*.

The genus *Bifidobacterium* was recognized as an independent genus only in the eighth edition of Bergey's Manual of Determinative Bacteriology, included in the family of *Actinomycetaceae* of the order *Actinomycetales* (Rogosa, 1974). Using 16S rRNA analysis, Stackebrandt and colleagues (1997) proposed a novel hierarchical structure in which the genera *Bifidobacterium* and *Gardnerella* were combined into the single family of *Bifidobacteriaceae* in the order of *Bifidobacteriales*, class *Actinobacteria*, phylum *Firmicutes* of the domain *Bacteria*.

The characteristic metabolic pathway of hexose fermentation in the *Bifidobacterium* genus was elucidated by Scardovi and Trovatelli (1965) and De Vries and colleagues (1967). Glucose is degraded exclusively by the fructose 6-phosphate shunt in which fructose-6-phosphoketolase cleaves fructose 6-phosphate into acetylphosphate and erythrose 4-phosphate. This fructose-6-phosphate phosphoketolase is a key enzyme of glycolytic fermentation in *Bifidobacterium* and serves as an identifying characteristic for the genus.

Currently, bifidobacteria are considered Gram-positive, anaerobic, non-motile, non-spore forming rod-shaped bacteria, with varied branching. They belong to the *Bifidobacteriaceae* family and have high genomic G+C content (55-67 mol%) (Biavati and Mattarelli, 2012). Bifidobacteria are known to colonize the human vagina, oral cavity, breast milk, and, more abundantly, the gastrointestinal tract (GIT) (Gomez-Gallego *et al.*, 2016; Biavati and Mattarelli, 2006). Beyond the human microbiome, they can be found in sewage, fermented milk products, and the gastrointestinal tracts of several animals including cattle/calf, chicken, lamb, rabbit, piglet, rat, and honeybees (Biavati and Mattarelli, 2006). Although members of the genus inhabit a wide range of habitats, most *Bifidobacterium* spp. are host-specific (Biavati and Mattarelli, 2006).

Significance of lactic acid bacteria

Lactic acid bacteria (LAB) are defined as a group of Gram-positive, catalase-negative, acid-tolerant, bacterial species that produce lactic acid as the main end-product of fermentation. The most common LAB are the rod-shaped bifidobacteria and lactobacilli, but LAB also comprises cocci like *Streptococcus, Pediococcus*, and *Leuconostoc* (Kandler, 1983). The first indication that high numbers of bifidobacteria and other lactic acid bacteria in the adult large intestine provides good health was proposed by Metchnikoff in 1908. He observed that aging adults had lower bifidobacteria counts than younger adults and that extended longevity of the Balkan people could be attributable to their habit of consuming fermented milk products like yoghurt (O'Sullivan and Kullen, 1998). Several other studies followed Metchnikoff's theory to investigate the potential of lactic acid bacteria consumption to improve human and animal health. This concept is now referred as 'probiotics' and is defined as 'live microorganisms which when administered in adequate amounts confer a health benefit on the host' (FAO and WHO, 2002). Probiotics are often incorporated in yogurt and fermented milk, but other dairy products like buttermilk, sour cream,

powdered milk, and frozen desserts, are also utilized for formulations. Production and consumption of fermented foods dates back many thousands of years, in which food is modified by the action of microorganisms or their products to achieve a desired change in the food (Naidu et al., 1999).

Since Tissier, bifidobacteria have been the subject of numerous studies due to their probiotic potential and health promoting characteristics, such as immune modulation (Fanning et al., 2012; Hart et al., 2004), production of bacteriocins (Martinez et al., 2013), improved gut metabolism (Sugahara et al., 2015) and inhibition of pathogens (Bernet et al., 1993; Arboleya et al., 2011; Fukuda et al., 2011). Several mechanisms of action of probiotics on intestinal epithelial cells have been proposed, including blocking pathogen entry into the epithelial cell by providing a physical barrier (colonization resistance); creating a mucus barrier by causing the release of mucus from goblet cells; maintaining intestinal permeability by increasing the intercellular integrity of apical tight junctions; producing antimicrobial factors; stimulating the innate immune system by signalling dendritic cells, which then travel to mesenteric lymph nodes and lead to the induction of regulatory T cells, and the production of anti-inflammatory cytokines, including IL-10 and TGF-β; preventing or triggering an innate immune response by initiating TNF production by epithelial cells and inhibiting (or activating) NFκB in macrophages and dampening (or priming) the host immune response by influencing the production of IL-8 and subsequent recruitment of neutrophils to sites of intestinal injury (Gareau et al., 2010).

Beyond the gastro-intestinal tract, lactic acid bacteria, such as *Lactobacillus* spp., are well-recognized organisms that promote good health in the female genital tract. Lactobacilli compete with urogenital pathogens (Boris *et al.*, 1998; Osset *et al.*, 2001; Phukan *et al.*, 2013; Atassi *et al.*, 2006) and produce bacteriocins, biosurfactants and hydrogen peroxide, that have all been shown to be microbicidal against several genital pathogens (Sabia et al., 2014; Stoyancheva et al., 2014; Ruíz et al., 2012; Satpute et al., 2016; Walencka et al., 2008). Nevertheless, production of lactic acid seems to be the main health-promoting feature of lactobacilli (O'Hanlon et al., 2013; Conti et al., 2009; Juárez Tomás et al., 2003; Aldunate et al., 2013; Boskey et al., 1999, 2001). The health-promoting features of other LAB like bifidobacteria in the vaginal microbiome have not yet been adequately investigated.

Mother to infant bifidobacteria transmission

While it is well known there is a microbial community residing in the human vagina, the upper genital tract, including uterus and fallopian tubes, is generally thought to be sterile. This would require the cervix to be a perfect barrier, but changes in the cervical mucus due to certain conditions may in theory allow bacteria to ascend to the uterus (Chen et al., 2017). A recent study demonstrated the existence of distinct bacterial communities throughout the female genital tract forming a continuum of microbial communities from the vagina to the ovaries (Chen et al., 2017). The placenta has also been reported to harbour its own microbiota (Aagaard et al., 2014). Although the potential of contamination during sample collection has raised some concerns regarding the conclusions from these studies (Lauder et al., 2016), their proposition that the upper genital tract, including the placenta, is not sterile challenges the traditional view of human fetal development as a sterile event. Nevertheless, in absence of infection, it is still generally accepted that the fetus is maintained in a sterile state, referred as the 'sterile womb paradigm', in which microbes are acquired both vertically (from the mother) and horizontally (from other humans or the environment) only during and after birth (Perez-Muñoz et al., 2017).

Hence, human microbial colonization likely begins at birth and continues to develop and change in composition for about 3 years, until the microbiota becomes adult-like (Arrieta et al., 2014). Early microbial colonization is an essential process for a healthy immune and metabolic maturation (Rautava et al., 2012). This initial colonization may be affected by many factors, such as the mode of delivery (vaginal or caesarean section), feeding type (breast-fed or formula-fed), exposure to antibiotics and hygiene (Fanaro et al., 2003; Mueller et al., 2015).

The newborn gut microbiota is generally initially colonized by enterobacteria that provide the conditions for anaerobe colonization (Arrieta et al., 2014). Around the first month of life, human milk oligosaccharides (HMOs) present in the maternal breast milk modulate neonatal gut microbial composition by promoting the growth of bifidobacteria (Mueller et al., 2015), which have also been shown to prevent gut infection (Fukuda et al., 2011; Penders et al., 2006). In this period, bifidobacteria may represent 60 to 91% of fecal bacteria in breast-fed infants (Harmsen et al., 2000; Turroni et al., 2012). This proportion decreases with age to less than 10% of the adult fecal microbiota (Yatsunenko et al., 2012; Turroni et al., 2008). The introduction of solid foods at around 4–6 months is accompanied by an increase of clostridial species; members of the

Ruminococcaceae family increase in abundance in the following months, and by 2–3 years of age, the microbiota composition consists of mainly *Bacteroidaceae*, *Lachnospiraceae*, and *Ruminococcaceae*, which then remains stable into adulthood (Arrieta et al., 2014).

Regardless of whether there is a microbial community residing in the upper genital tract, the contribution of maternal gut, vaginal and breast milk microbiota, specifically bifidobacteria, to infant microbiota is undeniable. Although the mere presence of common bacterial species does not necessarily indicate vertical transmission, numerous studies have been conducted to investigate and demonstrate the influence of maternal microbiotas on the neonatal gut microbiome (Matamoros et al., 2013; Bäckhed et al., 2015; Makino et al., 2013; Dominguez-Bello et al., 2010; Gabriel et al., 2017; Gomez-Gallego et al., 2016; Hunt et al., 2011; Milani et al., 2015; Duranti et al., 2017; Pannaraj et al., 2017; Asnicar et al., 2017; Takahashi et al., 2010; Sakwinska et al., 2017).

Asnicar and colleagues (2017) demonstrated vertical transmission of *B. bifidum* by applying shotgun metagenomic sequencing for strain-level profiling of fecal and breast milk samples from mother-infant pairs. Duranti and colleagues (2017) also demonstrated vertical transmission of bifidobacteria species and bifido(pro)phages from breast milk based on the internal transcribed spacer (ITS). In a study comparing mode of delivery, Bäckhed and colleagues (2015) observed that ~70% of the bacteria taxa present in the gut of vaginally delivered infants were also present in their mothers' gut microbiome; while in C-section infants vertical mother-infant transmission was less frequent for organisms like *Bifidobacterium*. Similarly, Domiguez-Bello and colleagues (2010) reported that a mother's vaginal bacterial community is significantly more similar to her own baby's microbiota than to the microbiota of other vaginally delivered babies, suggesting that the vaginal bacterial community is unique to each mother and is vertically transmitted to her baby. Millani and colleagues (2015) also reported vertical transmission of a *B. breve* and a *B. longum* subsp. *longum* strains from mother to infant by comparing mother's fecal and milk samples with the corresponding infant's fecal sample.

All these studies demonstrate and emphasize the importance of maternal microbiota collectively as a source of bacteria for the neonatal gut, especially bifidobacteria, which have several health-promoting characteristics in neonatal development. Nevertheless, the specific contribution of each bacterial community (gut, vaginal and milk) in the mother-to-infant

bifidobacteria transmission remains in discussion and more studies using high-resolution methods for comparison of strains are needed.

Genetic diversity of bifidobacteria

Considering the health-promoting characteristics of bifidobacteria in the gut, and consequently, their probiotic potential (Lee and O'Sullivan, 2010), there is a collective effort to increase the number of bifidobacteria genome sequences available. Genome sequence analysis provides new insights about bacterial evolution and helps to reveal the genetic attributes of bifidobacteria that contribute to their role in promoting gut health (O'Callaghan and van Sinderen, 2016). As a consequence of the emerging number of studies investigating the genomes of gut commensals, a new research theme has been proposed: the probiogenomics (Ventura et al., 2009a). The National Center for Biotechnology Information data base currently holds 424 publicly available bifidobacteria genome sequences, complete or draft, among which *B. longum*, *B. animalis*, *B. breve*, *B. bifidum*, and *B. adolescentis* are the species with the highest number of genome sequences available (https://www.ncbi.nlm.nih.gov/genome/?term=*Bifidobacterium*, November 2017).

The genus *Bifidobacterium* currently comprises ~50 species and according to taxonomic classification based on comparative analyses of 16S rRNA gene sequences and multilocus sequence typing based on several concatenated house-keeping genes (*clpC, dnaJ, xfp, dnaB, rpoC* and *purF*), bifidobacteria can be clustered into six phylogenetic groups: *B. boum, B. asteroides, B. adolescentis, B. pullorum, B. longum*, and *B. pseudolongum*. *B. asteroides*, isolated from the insect gut, appears to be the closest relative of the ancient progenitor of the genus *Bifidobacterium* (Bottacini et al., 2014; Ventura et al., 2006).

The average size of a *Bifidobacterium* genome is 2.2 Mb, with significant variation within the genus ranging from 1.7 (*B. indicum* and *B. coryneforme*) to ~3.3 Mb (*B. scardovi* and *B. biavati*). Such size differences suggest that bifidobacteria genomes have evolved as a result of many gene loss and/or acquisition events (Milani et al., 2014; Koskiniemi et al., 2012). *Bifidobacterium* is known to have high G+C content (mol%), with species *B. tsurumiense* (52.8%), *B. kashiwanohense* (56%), *B. saguini* (56.3%), *B. catenulatum* (56.4%) and *B. pseudocatenulatum* (56.4%) having the lowest G+C content, and species *B. gallinarum* (64.2%), *B. pullorum* (64.3%),

B. scardovi (64.6%), *B. cuniculi* (64.8%) and *B. choerinum* (65.7%) having the highest G+C content (Sun et al., 2015; Milani et al., 2014). Bifidobacteria typically encode 52–58 tRNA genes per genome, with the exception of *B. longum* subsp. *infantis* ATCC 15697 that has 79 tRNA-encoding genes (O'Callaghan and van Sinderen, 2016).

Comparison of *Bifidobacterium* spp. genome sequences has revealed a high degree of conservation and synteny across their genomes (Ventura et al., 2007). Nevertheless, phenotypic differences have been described among bifidobacteria, suggesting that niche adaptation has led to some degree of specialization (Ventura et al., 2009a). The human genome encodes 97 glycoside hydrolases, but just eight of those enzymes are directly linked to carbohydrate digestion, indicating that many complex dietary carbohydrates reach the colon undigested and therefore can be metabolized by (bifido)bacteria (Kaoutari et al., 2013). The types of carbohydrates in the diet have a significant impact on the metabolism of different groups of bacteria. While some members of the gut microbiota can use different substrates, from diet or host origin for example, others are more specialized. Bifidobacteria are presumably able to metabolize complex sugars derived from the diet as well as from the host, which gives them an ecological advantage among other members of the gut microbiota (Ventura et al., 2009a).

Analysis of bifidobacteria genomes reveals that more than 12% of the annotated open reading frames within their genome sequences are attributed to carbohydrate metabolism (Milani et al., 2014) and nearly 10% of the gene content in *B. longum* is dedicated to sugar internalization, via ABC transporters (Schell et al., 2002). Other members of the gut microbiota, such as *Escherichia coli* or *Enterococcus faecium*, and non-gut resident bacteria, such as *Lactococcus lactis*, have a substantially lower percentage of genes related to carbohydrate metabolism in comparison with bifidobacteria (Ventura et al., 2009a).

Genome analysis of *B. bifidum* and *B. longum* subsp. *infantis* revealed adaptations for the metabolism of human milk oligosaccharides (HMO) (Sela et al., 2008; Duranti et al., 2015). While *B. breve* and *B. longum* subsp. *longum* do not have the same efficient genetic repertoire for HMO degradation, they can utilize some HMOs and use carbohydrates that are broken down by other bifido(bacteria) via cross-feeding (Egan et al., 2014; Chaplin et al., 2015). In addition, *B. adolescentis* and *B. longum* subsp. *longum* can utilize plant-derived sugars, which are likely to be more abundant as diet changes after weaning. *B. bifidum* might also shift its HMO-metabolic

abilities towards mucin degradation when infants start weaning (Schell et al., 2002; Turroni et al., 2010; Sela et al., 2008; Duranti et al., 2014). Notably, genome analysis of *B. dentium* indicated adaptations towards the human oral cavity that include specialized nutrient acquisition and defences against antimicrobials (Ventura et al., 2009b).

Although the number of bifidobacteria genome sequences has been increasing over the years, most sequences available are from the species that commonly reside in the human gut, likely due their potential probiotic properties. Based on genomic studies of bifidobacteria, which often focus on carbohydrate metabolism, the genetic diversity within the *Bifidobacterium* genus is mostly related to the ability to metabolize different types of carbohydrates. Overall, the greater number of genes related to carbohydrate metabolism in bifidobacteria is an ecological advantage for this group in the gut microbiota and might indicate niche specialization for some bifidobacteria species. Lastly, genomic information is still very limited for other bifidobacteria species that are less common in the gut, and further studies are needed to better understand the genetic diversity of the whole *Bifidobacterium* genus.

OBJECTIVES

1. Describe the vaginal microbiome of pregnant women at low risk for preterm birth by comparing their microbial profiles to those of healthy non-pregnant women.

2. Assess whether there are differences in the vaginal microbiota composition early in gestation of women who had spontaneous preterm birth (sPTB) or term delivery that could be further investigated as diagnostic indicators of preterm birth risk.

3. Apply species-specific quantitative PCR to confirm the relative abundance of *Bifidobacterium* in the vaginal microbiomes of reproductive aged women previously determined based on high-throughput DNA sequencing methods, and characterize vaginal *Bifidobacterium* isolates based on carbohydrate fermentation patterns, hydrogen peroxide production, lactic acid production, resistance to low pH and lactic acid, and susceptibility to antibiotics.

4. Compare the genomes of gut and vaginal isolates of *B. breve* and *B. longum* to identify evidence of specialization that could indicate if vaginal and gut strains represent two distinct, niche-adapted subpopulations.

CHAPTER 2. The vaginal microbiome of pregnant women is less rich and diverse, with lower prevalence of Mollicutes, compared to non-pregnant women

Citation

Freitas AC, Chaban B, Bocking A, Rocco M, Yang S, Hill JE, Money DM & The VOGUE Research Group (2017). The vaginal microbiome of pregnant women is less rich and diverse, with lower prevalence of Mollicutes, compared to non-pregnant women. *Scientific Reports*, 7:9212. doi.org/10.1038/s41598-017-07790-9

Authors contributions

Conceived the study: DM, AB, JEH and the other members of the VOGUE Research Group. Oversaw and contributed to data collection and analysis, and participated in manuscript review: DM, AB, and JEH. Generated the Mollicutes PCR, data analysis and manuscript writing: ACF. Generated cpn60 vaginal microbiome profiles and 16S rRNA qPCR: ACF and BC. Recruited study participants and assisted in the collection of samples and data collection: MR. Assisted in the collection and processing of samples as well as study design: SY.

Chapter transition

The vaginal microbiome has a profound impact in both women's and neonatal health, and although the knowledge in the field has advanced in the past decades, still little is known about the effect of pregnancy on the vaginal microbiota. In this study chapter I focused on establishing a baseline description of the vaginal microbiome of healthy pregnant women. My objective was to characterize the vaginal microbiome of healthy pregnant women and compare them to those of healthy non-pregnant women. The results showed that overall microbiome cannot predict pregnancy status, although pregnancy affects the microbiome by mainly increasing lactobacilli abundance and decreasing Mollicutes prevalence. The ultimate goal of this study chapter was to establish a suitable description of the microbiome of healthy pregnant women that could be used as the "control group" for forthcoming studies focusing on specific cohorts, such as women who experienced preterm birth.

2.1. Abstract

The vaginal microbiome plays an important role in maternal and neonatal health. Imbalances in this microbiota (dysbiosis) during pregnancy are associated with negative reproductive outcomes, such as pregnancy loss and preterm birth, but the underlying mechanisms remain poorly understood. Consequently, a comprehensive understanding of the baseline microbiome in healthy pregnancy is needed. We characterized the vaginal microbiomes of healthy pregnant women at 11-16 weeks of gestational age (n=182) and compared them to those of non-pregnant women (n=310). Profiles were created by pyrosequencing of the cpn60 universal target region. Microbiome profiles of pregnant women clustered into six Community State Types (CST): I, II, III, IVC, IVD and V. Overall microbiome profiles could not be distinguished based on pregnancy status. However, the vaginal microbiomes of women with healthy ongoing pregnancies had lower richness and diversity, lower prevalence of *Mycoplasma* and *Ureaplasma* and higher bacterial load when compared to non-pregnant women. *Lactobacillus* abundance was also greater in the microbiomes of pregnant women with *Lactobacillus*-dominated CSTs in comparison with non-pregnant women. This study provides further information regarding characteristics of the vaginal microbiome of low-risk pregnant women, providing a baseline for forthcoming studies investigating the diagnostic potential of the microbiome for prediction of adverse pregnancy outcomes.

Keywords: vaginal microbiome, cpn60, pregnancy, *Lactobacillus*, *Mycoplasma*, *Ureaplasma*, community state type.

2.2. Introduction

The complex microbial community present in the female lower genital tract is an important factor in a woman's reproductive health. Imbalances in this microbiota can lead to bacterial vaginosis (BV), one of the most common gynaecologic conditions in women of reproductive age throughout the world (Ryckman et al., 2009; Sobel, 2000; Brotman, 2011). At present, it is understood that the healthy vaginal microbiome is dominated by *Lactobacillus* species, while BV is characterized by a relatively low abundance of *Lactobacillus* spp. accompanied by polymicrobial anaerobic overgrowth, including species such as *Gardnerella vaginalis*, *Prevotella*

spp., *Bacteroides* spp., *Mobiluncus* spp., and *Mycoplasma hominis* (Oakley et al., 2008; Zhou et al., 2010).

In pregnancy, such imbalances in the vaginal microbiome are associated with an increased risk of post-abortal infection (Larsson et al., 1992, 2000), early (Donders et al., 2000; Ralph et al., 2015) and late miscarriage (Llahi-Camp et al., 1996; Hay et al., 1994), histological chorioamnionitis (Gibbs, 1993; Hillier et al., 1988), postpartum endometritis (Watts et al., 1990; Jacobsson et al., 2002), preterm premature rupture of membranes (PPROM) (Parry and Strauss, 1998; Hillier et al., 1995) and preterm birth (Leitich et al., 2003). The relationship between BV and preterm birth in particular has profound implications since children who are born prematurely have higher rates of cardiovascular disorders, respiratory distress syndrome, neurodevelopmental disabilities and learning difficulties compared with children born at term (Behrman and Butler, 2007), along with increased risk of chronic disease in adulthood (Mwaniki et al., 2012). Also, preterm birth complications are estimated to be responsible for 35% of the world's annual neonatal deaths (Liu et al., 2012).

The development of culture-independent techniques, such as high throughput DNA sequencing, has greatly facilitated the comprehension of the composition and role of the vaginal microbial community. These tools can be used to exploit the potential of the microbiome to diagnose BV-associated conditions. However, while most studies have focused on the vaginal microbiome of healthy, non-pregnant, reproductive aged women, relatively little is known about this microbiome in pregnancy. Pregnancy is associated with a variety of physiological events including increased sex steroid hormone levels (Tulchinsky et al., 1972), host immune response modulation (Jamieson et al., 2006; Mor and Cardenas, 2010), altered immune-physicochemical properties of the cervical mucus (House et al., 2009; Luo et al., 2000; Lee et al., 2011a), as well as behavioural changes such as reduced drinking and smoking (Crozier et al., 2009). These factors may drive changes in the structure and/or composition of the microbial community resulting in a microbiome that is different from that of non-pregnant women. Thus, the current definition of a healthy microbiome, which has been based on findings in non-pregnant women, cannot necessarily be extrapolated to the microbiome in pregnancy. Establishing baseline data for the vaginal microbiome during pregnancy is a crucial step in the process of developing tools to predict and prevent pregnancy complications such as preterm birth, which is in turn associated with infection and/or inflammation as well as significant infant mortality and morbidity.

A few recent studies have used an amplicon sequencing approach to analyze the vaginal microbiome in pregnancy, all of which have been based on amplification and sequencing of variable regions of the 16S rRNA gene. The results of these studies suggest that pregnancy has a marked effect on the vaginal microbiome, leading to greater stability, increased *Lactobacillus* proportional abundance and reduced richness and diversity relative to the vaginal microbiomes of non-pregnant women (Aagaard et al., 2012; Romero et al., 2014c; Walther-António et al., 2014; MacIntyre et al., 2015; DiGiulio et al., 2015).

The objectives of this study were to describe the vaginal microbiome of pregnant women at low risk for preterm birth and to compare their microbial profiles to those of healthy non-pregnant women (Albert et al., 2015). This approach enabled direct comparisons between the two cohorts in which the main difference was their pregnancy status. Microbiome profiling was based on sequencing of the cpn60 universal target, which provides higher resolution than 16S rRNA variable regions (Links et al., 2012) and has allowed us to resolve previously undescribed vaginal microbiome community state types (Albert et al., 2015).

2.3. Methods

Study population and sampling

This study received ethical approval from the Mount Sinai Hospital Research Ethics Board (Approval Number 08-0005-A). All participants provided written informed consent and all methods were performed in accordance with the relevant guidelines and regulations. Women attending antenatal clinics at Mount Sinai Hospital (Toronto, ON, Canada) between May 2012 and October 2013 were invited to be part of a clinical trial to determine the effect of oral probiotic lactobacilli in altering the vaginal microbiome in asymptomatic pregnant women with an abnormal Nugent score (Nugent et al., 1991; Yang et al., 2015). Nugent score was determined on Gram-stained swabs taken at the same time as the swab for microbial analysis. Women with normal Nugent scores were excluded from the probiotic trial and are included in this analysis. Samples from women with an abnormal Nugent score who were subsequently randomized and included in the analysis for this study were taken prior to any intervention in these women and there was no difference in the pregnancy outcomes between the lactobacilli and placebo groups (Yang et al.,

2015). A preliminary report on the results of that probiotic trial have been presented previously (Yang et al., 2015).

Women were eligible to participate if the following inclusion criteria were met: currently pregnant, adequate comprehension of English language to sign written informed consent, age ≥18 years old, no evidence of fetal complications such as intrauterine growth restriction, and no evidence of medical complications of pregnancy. Exclusion criteria included inability to provide informed written consent, multi-fetal pregnancies, currently taking antibiotics or other antimicrobial therapy for BV treatment. Study data were collected and managed using REDCap electronic data capture tools (Harris et al., 2009).

Vaginal swabs were collected under direct visualization using a speculum by either a physician or a nurse and placed in dry tubes prior to being placed in -80˚C. A total of 182 pregnant women at 11-16 weeks gestation were enrolled in the vaginal microbiome study, including 111 women with normal Nugent scores (inconsistent with BV), 61 women with Nugent scores that were intermediate or consistent with BV, and 10 women with indeterminate Nugent scores due to poor quality smears. Total nucleic acid was extracted from swabs using the MagMAX™ Total Nucleic Acid Isolation Kit (Life Technologies, Burlington, ON, Canada) as per manufacturer's instructions. Kit reagents are aliquoted to eliminate repeated accessing of open reagents, and samples are processed in small batches using filter-tips to prevent cross-contamination. Pipettes and other lab surfaces are regularly treated with DNA surface decontaminant (DNA Away, ThermoFisher Scientific, Waltham, MA). Regular monitoring of reagent only DNA extraction controls in our lab by universal PCR confirms that these procedures are sufficient to eliminate detectable template contamination of study samples.

The microbial profiles of low-risk pregnant women were compared to profiles previously generated from healthy, reproductive aged, non-pregnant Canadian women from the greater Vancouver area, British Columbia, Canada (n=310) (Albert et al., 2015). Samples were collected as being non-menstrual but were not sampled at any specific non-menstrual cycle time as other studies have demonstrated there is little variation in microbiome profiles through the cycle (Chaban et al., 2014). Samples from this previous study were processed in the same way as in the current work in terms of swab type, storage temperature, DNA extraction, library preparation and

sequencing. Although the year of sampling was different between the two cohorts, there was no difference in time from sample collection to sequencing between the two groups.

Total Bacterial DNA (qPCR) and Detection of Mollicutes (PCR)

Quantitative PCR (qPCR): Total bacterial DNA quantity in each sample was estimated using a SYBR Green assay based on amplification of the V3 region of the 16S rRNA gene. Primer sequences were as follows: SRV3-1 (5'-CGGYCCAGACTCCTAC-3'), SRV3-2 (5'-TTACCGCGGCTGCTGGCAC-3') (Lee et al., 1996). Reactions run on a MyiQ thermocycler using the following cycling parameters: 95°C for 3 min, followed by 30 cycles of 95°C for 15 sec., 62°C for 15 sec., 72°C for 15 sec., with a final extension at 72°C for 5 minutes (Chaban et al., 2013).

Conventional PCR: Some Mollicutes (*Mycoplasma* and *Ureaplasma*) species lack a cpn60 gene (Hill et al., 2004). Thus, we performed a family-specific semi-nested PCR targeting the 16S rRNA gene to detect Mollicutes (van Kuppeveld et al., 1992), and a PCR targeting the multiple banded antigen gene to detect *Ureaplasma* spp.. In this assay, PCR products from *Ureaplasma parvum* and *U. urealyticum* can be differentiated by size (Watson et al., 1990).

cpn60 Universal Target (UT) PCR and Pyrosequencing

Universal primer PCR targeting the 552-558 bp cpn60 UT region was performed using a mixture of cpn60 primers consisting of a 1:3 molar ratio of primers H279/H280:H1612/H1613, as described previously (Chaban et al., 2014; Hill et al., 2006b; Schellenberg et al., 2011). To avoid cross-contamination, samples were handled in small batches, and a no template control was included with each set of PCR reactions. To allow multiplexing of samples in a single sequencing run, primers were modified at the 5' end with one of 24 unique decamer multiplexing identification (MID) sequences, as per the manufacturer's recommendations (Roche, Brandford, CT, USA). Amplicons were pooled in equimolar amounts for sequencing on the Roche GS Junior sequencing platform. The sequencing libraries were prepared using the GS DNA library preparation kit and emulsion PCR (emPCR) was performed with a GS emPCR kit (Roche Diagnostics, Laval, Canada).

Analysis of Operational Taxonomic Units (OTU)

Raw sequence data was processed by using the default on-rig procedures from 454/Roche. Filter-passing reads were used in the subsequent analyses for each of the pyrosequencing libraries. MID-partitioned sequences were mapped using Bowtie 2 (http://bowtie-bio.sourceforge.net/bowtie2/) on to a manually curated reference set of 1,561 OTU sequences representing human vaginal microbiota. Bowtie 2 was run using the default end-to-end alignment mode, in which the minimum "cutoff" for any individual read to be validly aligned to a reference sequence is an alignment score of -0.6 + -0.6 * L, where L is the length of the read. The best valid alignment for each read is reported. Mapping quality was also evaluated by MAPQ value, which is based on the probability that alignment does not correspond to the read's true point of origin.

The OTU reference set was generated originally by de novo assembly of cpn60 sequence reads from each of 546 vaginal microbiomes, which included 182 samples from pregnant women (this study) and 364 samples from non-pregnant women from previous studies by our research group. The reference assembly was created by the microbial Profiling Using Metagenomic Assembly pipeline (mPUMA, http://mpuma.sourceforge.net) (Links et al., 2013) with Trinity as the assembly tool (Grabherr et al., 2011). Assembled OTU were labeled according to their nearest reference sequence determined by watered-Blast comparison (Schellenberg et al., 2009) to the cpn60 reference database, cpnDB_nr (downloaded from http://www.cpndb.ca (Hill et al., 2004)). cpnDB_nr is a subset of the cpnDB database that includes a non-redundant collection of sequences representing all species in cpnDB, with a preference for inclusion of the type strain for each species when available. This reference assembly approach allows us to compare the microbial profiles from various cohorts under investigation, including the 182 pregnant women described in this study. To improve comprehension of some figures, we have pooled reads from OTU into "nearest neighbour species" based on their taxonomic label. Thus, the term "species" refers to OTUs that have the same nearest neighbour match in cpnDB.

Statistical Analysis

Comparisons across pregnancy status cohorts were based on analysis of variance (ANOVA), t-test and Chi-square, performed in IBM SPSS (Statistical Package for the Social Sciences, version 21) at 5% level of significance. For analysis of associations between

microbiological-socio-demographic characteristics and microbiome profiles, a false discovery rate (FDR) correction for multiple comparisons was applied (Benjamini and Hochberg, 1995).

Alpha (Shannon diversity and Chao1 estimated species richness) and beta diversity (jackknifed Bray-Curtis dissimilarity matrices) were calculated as the mean of 100 subsamplings of 1000 reads (or all reads available when less than 1000) in QIIME (Quantitative Insights Into Microbial Ecology) (Caporaso et al., 2010). Plots of alpha diversity measures against bootstrap sample number were generated in R and visually inspected to ensure that an adequate sampling depth for each sample was achieved. Microbiome profiles were also compared based on Bray-Curtis dissimilarity matrices using Principal coordinates analysis (PCoA) in QIIME.

For community state type analysis, a Jensen-Shannon distance matrix was calculated in R with a custom distance function that calculates the square root of the Jensen-Shannon divergence (Endres and Schindelin, 2003). This distance matrix was used for hierarchical clustering using the 'hclust' function in R, with Ward linkage.

Data availability

Raw sequence data files for the 182 samples described in this study were deposited to the NCBI Sequence Read Archive (Accession SRP073152, BioProject PRJNA317763). Due to ethical and legal restrictions related to protecting participant privacy imposed by the Mt. Sinai Hospital Research Ethics Board, all other relevant data are available upon request pending ethical approval. Please submit all requests to initiate the data access process to the corresponding author.

Comparison between microbiological and demographic characteristics of the pregnant cohort.

Microbiological and demographic characteristics were compared to CST (I, II, III, IVC, IVD, V), presence of Mollicutes (yes/no) and *Ureaplasma* (yes/no), microbiome richness (continuous variable) and diversity (continuous variable). Each of these five observations were compared to the following metadata: Nugent score category (BV-, intermediate BV, BV+), 16S RNA gene copy number (<104, 105-106, 107-108, >109), maternal age (18-25, 26-35, 36-45), maternal age (18-35, 35-45), BMI category (underweight, normal, overweight, obese), BMI (<25,

≥25), pre-existing conditions (yes/no), surgeries in the past 10 years (yes/no), smoking (yes/no), drinking alcohol (yes/no), antibiotics at enrolment (yes/no), gestational age at delivery (<37, ≥37), mode of delivery (vaginal, C-section), mode of delivery (vaginal, C-section, C-section-elective), fetal sex (male, female), neonate in level 2 care nursery (yes/no), birth weight (continuous variable) and Apgar score at 5 minutes (0-9).

2.4. Results

Cohort description

Socio-demographic characteristics were comparable between pregnant and non-pregnant participants in terms of age, BMI, ethnicity and smoking status (Table 2.1). There was no significant difference in BMI (t-test, p=0.211) or ethnicity (Chi-square, p=0.372). However, pregnant women (age 33 ± 4) were on average older than non-pregnant women (age 31 ± 7) (t-test, p<0.0001). The number of current smokers also differed between pregnant (2%) and non-pregnant women (12%) (t-test, p<0.0001).

Sixty-four percent of women described themselves as of Caucasian (white) background in the pregnant cohort, which was not different from the non-pregnant cohort. This was followed by East Asian (14%) and South Asian (7%). Ten percent of participants had other or mixed ethnicity. Pregnancy outcome data was not available for five women (5/182) since they delivered in different hospitals. Sixty women (33%) were nulliparous and the average gestational ages at enrolment and at delivery were 13^{+2} and 39^{+2} weeks, respectively (Table 2.2). Seventeen women (9.3%) had assisted conception and the majority of women had a vaginal delivery (74%) as opposed to C-section (26%). All infants were liveborn and seven were delivered prior to 37 weeks gestation. Average birth weight was 3376 g and only three newborns weighed less than 2500 g at birth, with one of these being preterm (<37 weeks gestation). Vaginal microbiomes from the women who delivered infants less than 2500 g were determined to belong to CST II (*L. gasseri* dominated) (1/3) and III (*L. iners* dominated) (2/3). Five infants were admitted to a level 3 intensive-care unit (NICU), four of which were preterm. None of these five infants had low birth weight. Vaginal microbiomes of the mothers of the five infants were determined to belong to CST III (*L. iners* dominated) (1/5), V (*L. jensenii* dominated) (2/5), IVC (*G. vaginalis* subgroup A dominated) (1/5) and IVD (mixed microbiome) (1/5).

Table 2.1. Sociodemographic, clinical and microbiological characteristics across pregnant and non-pregnant cohort.

[1]t-test; [2]Chi-square. *Characteristic with significant differences (at 5% level) between pregnant and non-pregnant cohorts. [+] [(total copies of 16S rRNA gene) / swab]

Characteristics	Pregnant	Non-pregnant
Age [1] *	(n=182)	(n=310)
(Mean ± SD, Range)	33.6 ± 4.3 (21 - 45)	30.1 ± 7.6 (18.6 - 49.2)
18-25	5 (2.7%)	111 (36%)
26-35	122 (67%)	131 (42%)
36-49	55 (30.2%)	68 (22%)
Body mass index [1]	(n=182)	(n=307)
(Mean ± SD, Range)	23.1 ± 4.21 (17 - 43)	23.9 ± 5.2 (15 - 50)
Underweight (<18.50)	7 (3.8%)	18 (5.8%)
Normal weight (18.51-24.9)	140 (76.9%)	194 (62.6%)
Overweight (25.0-29.9)	26 (14.3%)	55 (17.7%)
Obese (>30)	9 (4.9%)	40 (12.9%)
Ethnicity [2]	(n=182)	(n=306)
White	117 (64.3%)	200 (64.5%)
East Asian	26 (14.3%)	60 (19.4%)
South Asian	13 (7.1%)	12 (3.9%)
Black	4 (2.2%)	9 (2.9%)
Other/Mixed ethnicity	22 (12%)	25 (8%)
Status	(n=182)	(n=310)
Smoking [2] *	4 (2.2%)	39 (12.6%)
Alcohol [2]	11 (6%)	no data
Unprotected sex [2] *	(in the past 4 days)	(in the past 2 days)
Yes	47 (25.8%)	45 (14.5%)
Nugent category [2] *	(n=172)	(n=307)
Not consistent with BV	111 (61%)	250 (80.6%)
Intermediate BV	36 (20%)	25 (8%)
Consistent with BV	25 (14%)	32 (10.3%)
CST [2] *	(n=182)	(n=310)
I	56 (30.7%)	156 (50.3%)
II	12 (6.6%)	0 (0%)
III	30 (16.5%)	50 (16.1%)
IVA	0 (0%)	36 (11.6%)
IVC	26 (14.3%)	22 (7.1%)
IVD	42 (23.0%)	24 (7.7%)
V	16 (8.8%)	22 (7.1%)
[+]**Estimated bacterial load** [1] *	(n=181)	(n=309)
(Mean ± SD, Range)	7.77 ± 0.93 (4.89 - 10.67)	6.83 ± 1.55 (3.50 - 10.31)
10^4 or less	1 (0.5%)	47 (15.2%)
10^5-10^6	27 (15%)	97 (31.4%)
10^7-10^8	133 (73.5%)	146 (47.2%)
10^9 or more	20 (11%)	19 (6.1%)
Presence of Mollicutes [2] *	74 (40%)	217 (70%)
Presence of *Ureaplasma* [2] *	42 (23%)	149 (48%)
U. parvum	39 (21.4%)	127 (41%)
U. urealyticum	3 (1.6%)	5 (1.6%)
Both	0 (0%)	1 (0.3%)

Table 2.2. Pregnant cohort description and pregnancy outcomes.

Characteristics	Descriptive
Folic acid	(n=182)
Before conception	49 (26.9%)
During pregnancy	50 (27.5%)
Vitamins	(n=182)
Before conception	100 (54.9%)
During pregnancy	174 (95.6%)
Natural conception	(n=182)
No	18 (9.9%)
Parity	(n=182)
0	60 (33%)
1	98 (53.8%)
2-4	24 (13.2%)
Pre existing medical condition	(n=182)
Hypo/Hyperthyroidism	16 (8.8%)
Depression	9 (4.9%)
Asthma	6 (3.3%)
Anemia	5 (2.7%)
Other condition	41 (22.5%)
No	105 (57.7%)
Surgeries in the past 10 years	(n=182)
Yes	83 (45.6%)
Antibiotics at enrolment (for other infections excluding vaginitis/BV)	(n=182)
Amoxicillin/Penicillin	8 (4.4%)
Nitrofurantoin	3 (1.6%)
Other	3 (1.6%)
Unknown/Unsure	6 (3.3%)
No	162 (89%)
Gestational age at enrollment	(n=182)
(Mean ± SD, Range)	$13^{+2} \pm 1^{+1}$ (11^{+1} - 16^{+6})
Gestational age at delivery	(n=177)
(Mean ± SD, Range)	$39^{+2} \pm 1^{+2}$ (32^{+1} - 41^{+2})
Preterm birth (< 37 weeks)	7 (4%)
Mode of delivery	(n=177)
Vaginal delivery	131 (74.0%)
C-section	46 (26%)
Fetal sex	(n=177)
Female (%) / Male (%)	91 (51%) / 86 (49%)
Birth weight (g)	(n=177)
(Mean ± SD, Range)	3376 ± 474 (1876 - 5200)
Low (< 2500 g)	3 (1.7%)
Normal (2500 - 4200 g)	166 (93.8%)
Large (> 4200 g)	8 (4.5%)
Apgar score at 5 min	(n=177)
(Mean ± SD, Range)	8.97 ± 0.17 (8 - 9)
Neonate in level 3 care nursery after birth	(n=176)
	5 (2.8%)

Sequencing results and OTU analysis

We characterized the vaginal microbiomes of pregnant women at low risk of preterm birth using pyrosequencing of the universal target region of the cpn60 gene. Sequence reads from the vaginal microbiomes of pregnant women were mapped on to a manually curated reference set of 1,561 OTU sequences as described in the methods. Raw sequence data files for the 182 samples described in this study were deposited to the NCBI Sequence Read Archive (BioProject PRJNA317763). A total of 1,415,117 cpn60 reads was generated. Median and average read count per sample was 5,024 and 7,775 (range 494 – 43,245), respectively. Average read length was 448 bp. The average MAPQ value was 21.1.

Results of Bowtie2 mapping showed that these reads corresponded to 645 OTUs from the reference assembly (Supplementary Table 2.4S). A total of 82 OTUs (corresponding to 53 nearest neighbour "species") were at least 1% abundant in at least one sample. And only 22 "species" were detected in at least 50% of samples (Table 2.3). Although the ranges of percent identity to reference sequences are large in some cases, reflecting the diversity in the community, the most abundant OTU were at the high end of the range. In fact, of the 25 most abundant OTU in the study (accounting for 95% of the sequence reads generated), 22/25 were >95% identical to their nearest neighbour (Supplementary Table 2.4S).

Most reads (68.8%) were identified as *Lactobacillus* spp. And only three OTUs, all of which matched to *Lactobacillus* spp., accounted for 55.7% of all reads generated: OTU 1403: *L. crispatus* (30.6%), OTU 1174: *L. iners* (15.4%) and OTU 1479: *L. jensenii* (9.7%). Species with the highest sample prevalence were *L. jensenii* (181/182), *Streptococcus devriesei* (180/182), *L. crispatus* (178/182), *L. acidophilus* (178/182), followed by *Atopobium vaginae* (176/182) and *Weissella viridescens* (173/182). Even though OTUs with similarity to *Streptococcus devriesei* and *Weissella viridescens* were detected in most samples, they had very low abundance, representing only 0.33% and 0.22% of all reads, respectively (Supplementary Table 2.4S).

Table 2.3. Prevalence and proportion of total reads for "species" detected in at least 50% of samples.

Nearest neighbour[a]	% OTU identity range	Prevalence /182 (%)	% total reads
Lactobacillus jensenii	81.6 - 100	181 (99.4)	10.66
Streptococcus devriesei	83	180 (98.9)	0.34
Lactobacillus crispatus	78.8 - 99.8	178 (97.8)	31.93
Lactobacillus acidophilus	91.8 - 100	178 (97.8)	1.22
Atopobium vaginae	82.5 - 96.9	176 (96.7)	1.62
Weissella viridescens	58.8 - 99.5	173 (95.0)	0.22
Desulfotalea psychrophila	76.1	167 (91.7)	0.12
Lactobacillus iners	86.8 - 100	165 (90.6)	15.85
Lactobacillus gasseri	65.4 - 100	164 (90.1)	5.69
Streptococcus parasanguinis	94.9 – 97.0	162 (89.0)	0.08
Prevotella tannerae	77.5 - 98.7	162 (89.0)	0.07
Faecalibacterium cf. prausnitzii	77.5 - 79.6	159 (87.3)	0.16
Gardnerella vaginalis subgroup C	78.5 - 99.3	156 (85.7)	4.45
Peptoniphilus harei	89.9 - 98.5	156 (85.7)	0.07
Clostridium innocuum	75.4	148 (81.3)	0.06
Sphingobium yanoikuyae	99.3 - 99.4	145 (79.6)	0.08
Eubacterium siraeum	84.4	144 (79.1)	0.05
Gardnerella vaginalis subgroup B	93.2 - 99.6	133 (73.0)	0.54
Massilia timonae	82	126 (69.2)	0.03
Gardnerella vaginalis subgroup A	86.7 - 99.6	123 (67.5)	7.64
Bifidobacterium breve	89.6 - 99.5	119 (65.3)	2.12
Megasphaera sp. genomsp. type 1	58.1 - 86.7	112 (61.5)	3.09

[a]Closest match in the cpnDB reference database based on sequence identity.

Community state type analysis

Hierarchical clustering of vaginal microbiome profiles from pregnant women resulted in the resolution of six Community State Types (CST) (Figure 2.1). All *Lactobacillus*-dominated CST, previously described by Ravel and Gajer (Ravel et al., 2011) based on pyrosequencing of the 16S rRNA gene, were detected. Most profiles (114/182) were dominated by one of four *Lactobacillus* species that define four different CST: CST I (*L. crispatus*, n=56), CST II (*L. gasseri*, n=12), CST III (*L. iners*, n=30), and CST V (*L. jensenii*, n=16). All non-*Lactobacillus*-dominated samples were assigned to CST IVC or CST IVD, as previously described by Albert et al. (Albert et al., 2015). CST IVC (n=26) is dominated by *Gardnerella vaginalis* subgroup A, *Megasphaera* sp. *genomosp* type 1, and *G. vaginalis* subgroup C. CST IVD (n=42) was the most heterogeneous group, including a mixture of *Bifidobacterium dentium*, *B. infantis*, *B. breve*, *L. delbrueckii, Alloscardovia omnicolens*, *G. vaginalis* subgroup C, and *Atopobium vaginae.*

CST distribution of the pregnant participants was compared to the previously described non-pregnant cohort. Twelve pregnant women had microbiome profiles identified as the *L. gasseri*-dominated CST II, which was not observed among the profiles of non-pregnant women (Albert et al., 2015). Additionally, CST IVA was not detected among pregnant women, including 61 with intermediate or high Nugent scores. In the non-pregnant group, 14/57 (24.6%) of women with intermediate or high Nugent scores were assigned to CST IVA, with the remainder in CST IVC (15/57, 26.3%), CST IVD (20/57, 35.1%) or one of the *Lactobacillus* dominated CST (8/57, 14.0%). CST IVA was defined by Albert et al. (Albert et al., 2015) as a very heterogeneous group, dominated by *G. vaginalis* subgroup B or *Atopobium vaginae*, or combinations of *Stapylococcus, Streptococcus, Prevotella, Alloscardovia*, *Gardnerella*, *Bifidobacterium* and *Lactobacillus*. As expected, *Lactobacillus*-dominated CST, i.e. CST I, CST II, CST III and CST V, were associated with low Nugent score samples (BV negative), while CST IVC and CST IVD were associated with intermediate and high Nugent scores (BV positive) (Figure 2.1). Although there were differences regarding presence/absence of specific CST in the pregnant and non-pregnant cohorts, overall microbial profiles could not be distinguished based on pregnancy status alone (Figure 2.2 A-B).

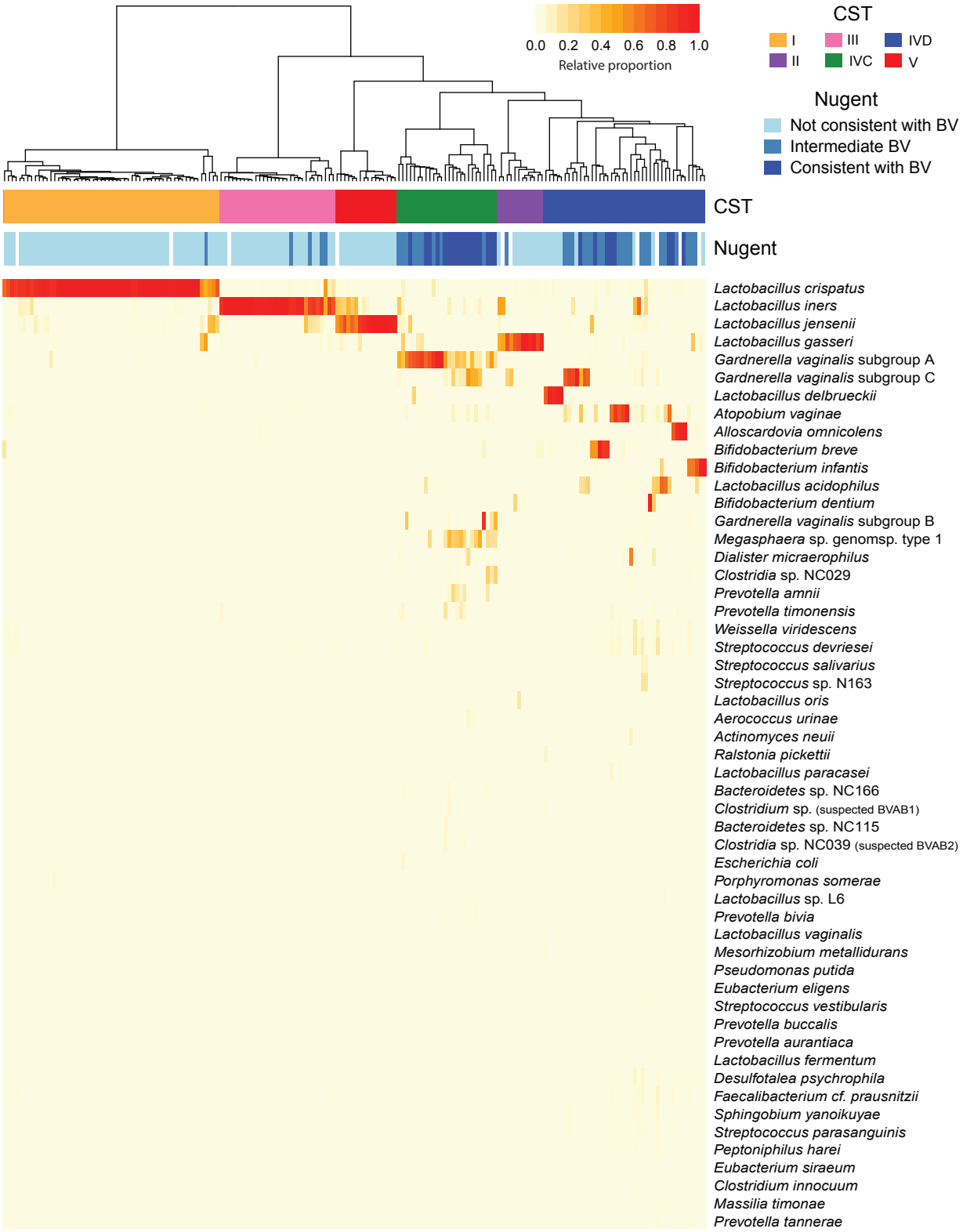
0.0 0.2 0.4 0.6 0.8 1.0
Relative proportion
CST
I
II
III
IVC
IVD
V
Nugent
Not consistent with BV
Intermediate BV
Consistent with BV
CST
Nugent
Lactobacillus crispatus
Lactobacillus iners
Lactobacillus jensenii
Lactobacillus gasseri
Gardnerella vaginalis subgroup A
Gardnerella vaginalis subgroup C
Lactobacillus delbrueckii
Atopobium vaginae
Alloscardovia omnicolens
Bifidobacterium breve
Bifidobacterium infantis
Lactobacillus acidophilus
Bifidobacterium dentium
Gardnerella vaginalis subgroup B
Megasphaera sp. genomsp. type 1
Dialister micraerophilus
Clostridia sp. NC029
Prevotella amnii
Prevotella timonensis
Weissella viridescens
Streptococcus devriesei
Streptococcus salivarius
Streptococcus sp. N163
Lactobacillus oris
Aerococcus urinae
Actinomyces neuii
Ralstonia pickettii
Lactobacillus paracasei
Bacteroidetes sp. NC166
Clostridium sp. (suspected BVAB1)
Bacteroidetes sp. NC115
Clostridia sp. NC039 (suspected BVAB2)
Escherichia coli
Porphyromonas somerae
Lactobacillus sp. L6
Prevotella bivia
Lactobacillus vaginalis
Mesorhizobium metallidurans
Pseudomonas putida
Eubacterium eligens
Streptococcus vestibularis
Prevotella buccalis
Prevotella aurantiaca
Lactobacillus fermentum
Desulfotalea psychrophila
Faecalibacterium cf. prausnitzii
Sphingobium yanoikuyae
Streptococcus parasanguinis
Peptoniphilus harei
Eubacterium siraeum
Clostridium innocuum
Massilia timonae
Prevotella tannerae

Figure 2.1. Vaginal microbial profiles of pregnant women at low risk of preterm birth.

Heatmap of hierarchical clustering of Jensen-Shannon distance matrices with Ward linkage on the relative proportions of reads for each OTU within individual vaginal samples (n=182). Each column represents a woman's vaginal microbiome profile, and each row represents an OTU. Only OTUs that are at least 1% abundant in at least one sample are shown. The proportion of the total microbiome comprised is indicated by white to red colour according to the legend. The coloured bars above the heatmap show the community state type (CST) and the Nugent score category (Nugent) for each woman. Legend: white = missing data.

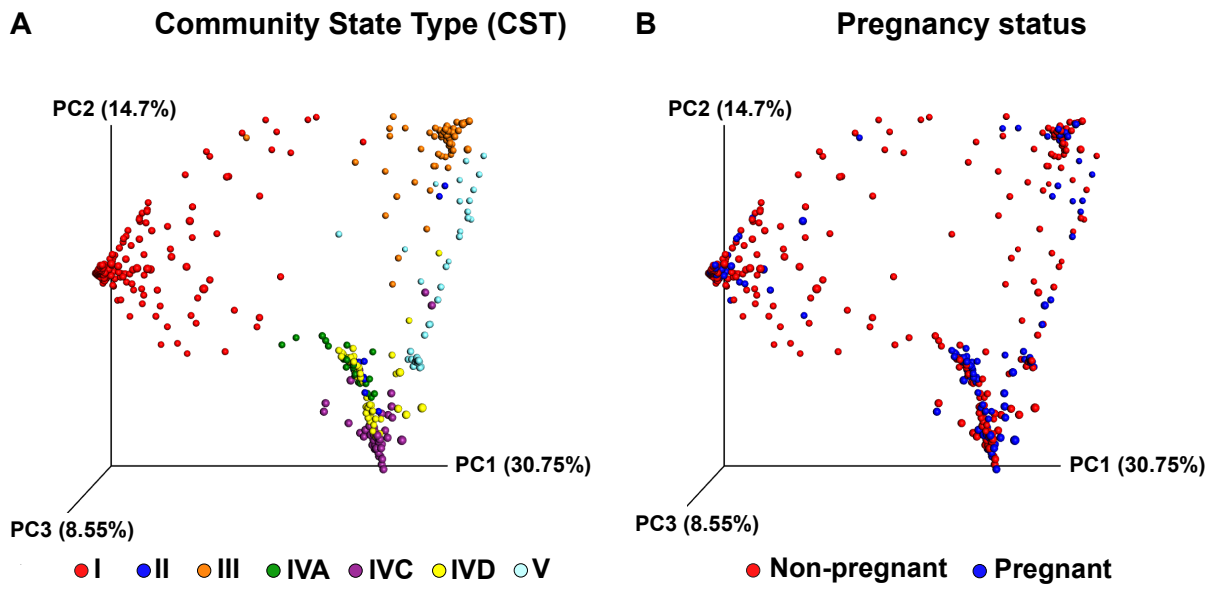

Figure 2.2. CST and pregnancy status of participants.

Jackknifed principal coordinates analysis (PCoA) of Bray-Curtis distance matrices of microbial profiles from all participants in the study, with individuals coloured by CST (A) or pregnancy status (B). Samples with fewer than 1000 sequence reads (16/492) were not plotted.

Abundance of *Lactobacillus* spp.

The proportion of *Lactobacillus* spp. in vaginal samples followed a bimodal distribution (Figure 2.3). Microbiomes of most pregnant women had either low (0-20%, n=46/182) or high (80-100%, n=113/182) *Lactobacillus* spp. abundance, whereas only 23/182 (12.6%) of samples had intermediate levels (20-80%) of *Lactobacillus* spp. Also, pregnant women in *Lactobacillus*-dominated CST had greater proportions of lactobacilli (95.7% ± 6.2) (i.e. were more *Lactobacillus* dominated) than non-pregnant women (89.9% ± 14.8) (t-test, $p<0.0001$) (Figure 2.4A).

Ecological analysis

Assessment of alpha diversity revealed that microbiomes of pregnant women were less diverse (Shannon diversity index, 1.3 ± 0.9) and less rich (Chao1, 38.9 ± 15.3) when compared to those of non-pregnant women (1.6 ± 1.1; 47.6 ± 26) (t-test, $p<0.001$) (Figure 2.4B-C). When comparisons were conducted within CST, profiles of pregnant women in CST I, III and V were less diverse and less rich than profiles of non-pregnant women in the same category (t-test, $p<0.05$) but no statistically significant differences were observed in CST IVC and IVD.

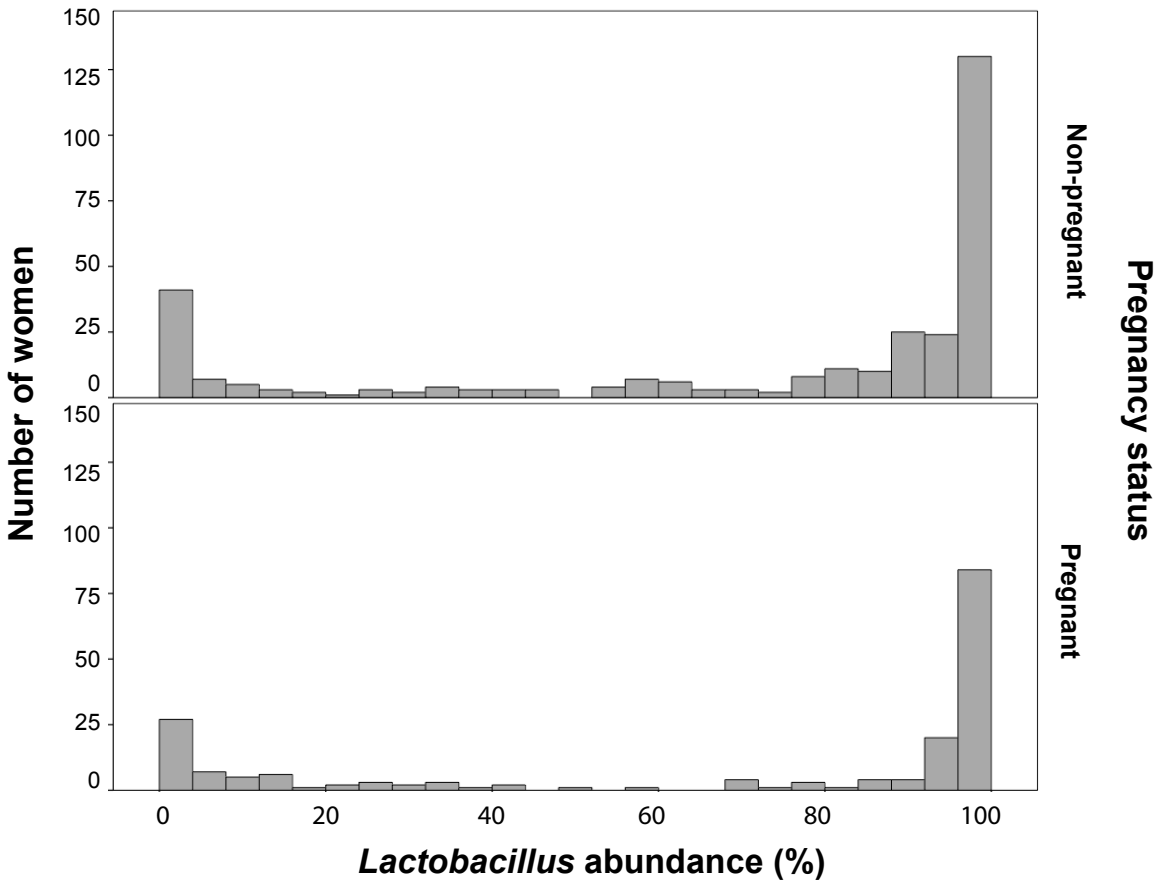

Figure 2.3. *Lactobacillus* spp. abundance.

Bimodal distribution of vaginal microbiome profiles of non-pregnant (upper panel) and pregnant (lower panel) women based on *Lactobacillus* spp. abundance (%).

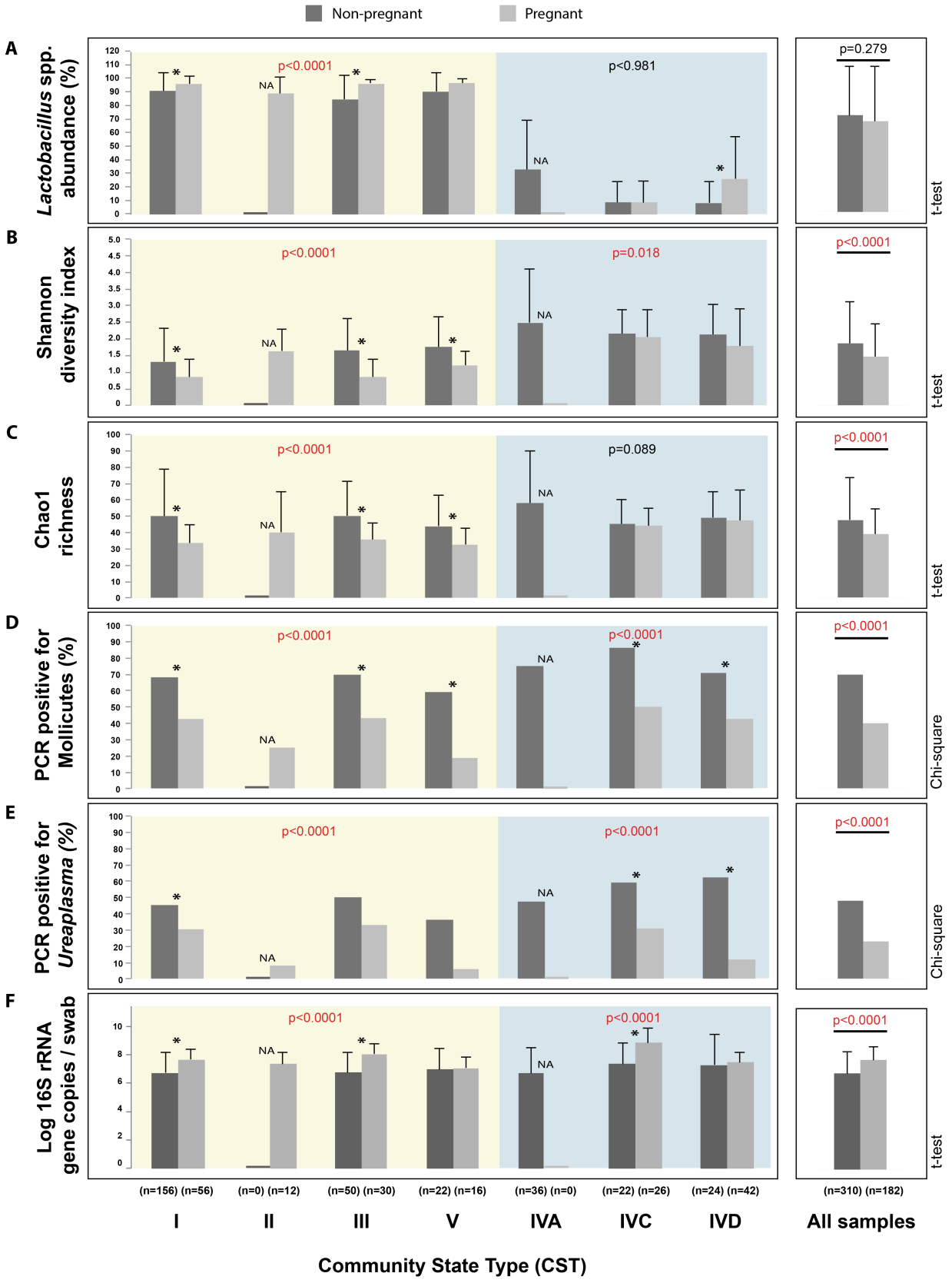
Non-pregnant
Pregnant
A
Lactobacillus spp. abundance (%)
p<0.0001
p<0.981
p=0.279
t-test
B
Shannon diversity index
p<0.0001
p=0.018
p<0.0001
t-test
C
Chao1 richness
p<0.0001
p=0.089
p<0.0001
t-test
D
PCR positive for Mollicutes (%)
p<0.0001
p<0.0001
p<0.0001
Chi-square
E
PCR positive for Ureaplasma (%)
p<0.0001
p<0.0001
p<0.0001
Chi-square
F
Log 16S rRNA gene copies / swab
p<0.0001
p<0.0001
p<0.0001
t-test
NA
(n=156) (n=56)
(n=0) (n=12)
(n=50) (n=30)
(n=22) (n=16)
(n=36) (n=0)
(n=22) (n=26)
(n=24) (n=42)
(n=310) (n=182)
I
II
III
V
IVA
IVC
IVD
All samples
Community State Type (CST)

Figure 2.4. Comparison of the microbial community features between pregnant and non-pregnant participants within each CST.

Lactobacillus spp. abundance (A), Shannon diversity (B), Chao1 richness (C), Mollicutes prevalence (D), *Ureaplasma* prevalence (E) and bacterial load (F) were compared between pregnant and non-pregnant women in each CST. For continuous variables (A, B, C, F), the mean value is plotted with error bars indicating standard deviation. Significant differences ($p<0.05$) between pregnant and non-pregnant women within each CST are indicated by an asterisk. p-values in the main panels refer to the comparison between pregnant and non-pregnant women in *Lactobacillus*-dominated CST (yellow panel) and non-*Lactobacillus*-dominated CST (blue panel). A comparison of pooled data from all CST is shown in the right-most panel. Statistical tests used are indicated on the right side of the graph.

Prevalence of Mollicutes and *Ureaplasma* (PCR)

Mollicutes (*Mycoplasma* and/or *Ureaplasma*) were detected by family-specific conventional PCR in 74/182 (40%) of pregnant women, but prevalence varied between CST: 43%, 25%, 43%, 18%, 50% and 43% in CST I, II, III, V, IVC and IVD, respectively (Figure 2.4D). *Ureaplasma* species were detected by genus-specific PCR in samples of 42/182 (23%) pregnant women, with 39 testing positive for *U. parvum* and 3 for *U. urealyticum* (Figure 2.4E). All pregnant women PCR positive for *U. urealyticum* (3/3) were in CST III. Pregnant women were less likely to be positive for Mollicutes detection by PCR when compared to non-pregnant women, regardless of CST (Chi-square, $p<0.0001$, Mollicutes positive: 217/310 non-pregnant women and 74/182 pregnant women) (Figure 2.4D). Additionally, pregnant women in both *Lactobacillus*- and non-*Lactobacillus*-dominated CST had lower prevalence of *Ureaplasma* species than non-pregnant women (Chi-square, $p<0.0001$, *Ureaplasma* positive: 149/310 non-pregnant women and 42/182 pregnant women) (Figure 2.4E).

Total bacterial load (qPCR)

Total bacterial load was assessed based on qPCR targeting the 16S rRNA gene, and expressed as log copy number per swab (Figure 2.4F). Average bacterial load was log 7.77 ± 0.93 16S rRNA gene copies per swab (range log 4.89 – 10.67). Bacterial load within each CST for the pregnant women's samples were: CST I (log 7.6 ± 0.7), CST II (log 7.3 ± 0.8), CST III (log 8.0 ± 0.7), CST V (log 7.0 ± 0.7), CST IVC (log 8.8 ± 1) and CST IVD (log 7.4 ± 0.7).

Higher bacterial loads were detected in samples from pregnant women (log 7.7 ± 0.9) when compared to non-pregnant women (log 6.8 ± 1.5) (t-test, $p<0.0001$) (Figure 2.4F). Comparisons within CST confirmed that samples from pregnant women in CST I (log 7.6 ± 0.7), CST III (log 8.0 ± 0.7) and CST IVC (log 8.8 ± 1) had higher bacterial load than non-pregnant women in the same categories (CST I = log 6.7 ± 1.4, CST III = log 6.7 ± 1.4. CST IVC = log 7.3 ± 1.5) (t-test, $p<0.0001$). In a second analysis, samples were pooled into two groups: *Lactobacillus*- (I, II, III, V) and non-*Lactobacillus* dominated (IVA, IVC, IVD) CST. Bacterial load values were statistically different (t-test, $p<0.0001$) between pregnant and non-pregnant women in both groups, with pregnant women having greater bacterial load (*Lactobacillus*-dominated CST: log 7.6 ± 0.8,

non-*Lactobacillus* dominated CST: log 8 ± 1) than non-pregnant women (*Lactobacillus*-dominated CST: log 6.7 ± 1.4, non-*Lactobacillus* dominated CST: log 7.0 ± 1.8).

Relationships between microbiological and socio-demographic characteristics across the pregnant cohort

The characteristics of the microbial community of pregnant women were analyzed in terms of their relationship with the socio-demographic and clinical data. First, we determined whether there was any relationship between the CST (I, II, III, IVC, IVD, V) and demographic characteristics such as BMI, ethnicity, unprotected sex, folic acid intake, vitamins, natural conception, antibiotic use, gestational age at delivery, mode of delivery, neonatal in high level care nursery, parity, pre-existing conditions, surgeries (past 10 years), smoking and alcohol drinking status as well as Nugent score (Figure 2.5). Besides Nugent score, the only significant interaction was between CST and parity (0 or ≥1) (Chi-square, p=0.033), with 45% of women at parity 0 in CST I (27/60) and 23% of women at parity ≥ 1 in CST I (29/122). Microbiological and demographic characteristics were also compared to presence of Mollicutes (yes/no) and *Ureaplasma* (yes/no), microbiome richness (continuous variable) and diversity (continuous variable). These four observations were compared to 29 other variables. PCR detection of Mollicutes (p=0.017) and *Ureaplasma* (p=0.017) was significantly associated with bacterial load (Chi-square).

Although smoking has been previously shown to have a possible effect on the vaginal microbiota (Brotman et al., 2014a), there was no significant association between CST and smoking status in this study (Chi-square, p>0.05). Also, microbiome analysis was redone excluding samples from participants who are smokers (results not shown). The results led to the same conclusions regarding the overall microbiome (PCoA), *Lactobacillus* abundance, Shannon diversity, Chao1 richness, bacterial load and Mollicutes/*Ureaplasma* prevalence.

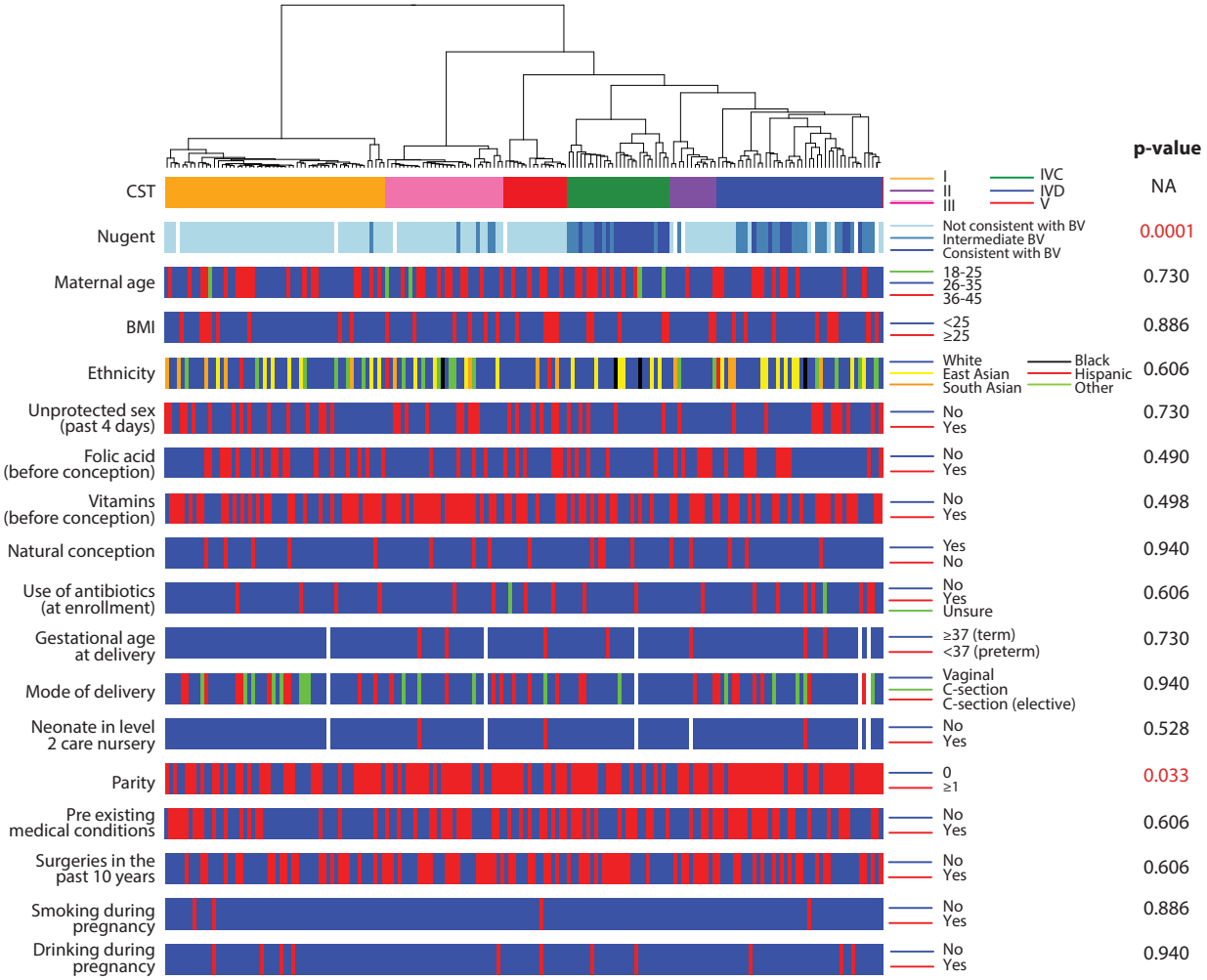

Figure 2.5. Socio-demographic characteristics of pregnant participants in relation to CST.

Hierarchical clustering of microbiome profiles based on Jensen-Shannon distance matrices with Ward linkage of the relative proportions each OTU within individual vaginal samples (n=182). Demographic characteristics are indicated on the left side, and categories indicated on the right side of each row. Numbers on the right side indicate adjusted p-values of Chi-square test after false discovery rate correction. Fisher's Exact test was conducted for variables where at least one category had an expected frequency of less than 5. Legend: white = missing data.

2.5. Discussion

In this study we characterized the vaginal microbiome of pregnant women with healthy ongoing pregnancies, at low risk of experiencing pregnancy complications such as preterm birth, and compared these results to our previously characterized cohort of non-pregnant women of similar ethnicity. We focused our studies on the vaginal microbiome at 11-16 weeks. Pregnancy pathologies such as spontaneous preterm birth and early onset pre-eclampsia have their origins in the first or early second trimester and therapeutic interventions at this stage have been shown to be efficacious (Roberge et al., 2012). This is also the gestational age at which pregnant women in Canada often have their first prenatal visit with a health care provider and a vaginal swab is taken to assess the presence or absence of vaginal/cervical infections.

We used the cpn60 universal target region in our study due to its superior resolution of bacterial taxa relative to 16S rRNA sequences (Paramel Jayaprakash et al., 2012; Vermette et al., 2010; Verbeke et al., 2011; Hill et al., 2006a; Lan et al., 2016; Masson et al., 2006; Brousseau et al., 2001), its establishment as a preferred barcode for bacteria (Links et al., 2012), and the availability of a reference database of vaginal bacterial cpn60 sequences from previous studies (Hill et al., 2005b; Schellenberg et al., 2011; Chaban et al., 2014; Albert et al., 2015; Schellenberg et al., 2012). This approach has been applied to numerous animal (Mckenney et al., 2014; Costa et al., 2014; Desai et al., 2012), human (Peterson et al., 2016) and environmental microbial populations (Bondici et al., 2013; Dumonceaux et al., 2006; Links et al., 2014; Oliver et al., 2008) and has previously resulted in discovery of novel community state types based on the resolution of distinct subgroups of *Gardnerella vaginalis* (Albert et al., 2015).

In order to make valid comparisons between pregnant and non-pregnant women microbiomes, we analyzed the results from both cohorts based on socio-demographic characteristics (Table 2.1). The two cohorts were comparable, with no significant differences detected in any category except for maternal age and smoking. Differences regarding smoking status were not surprising since behavioural changes such as reduced drinking and smoking have been documented in pregnancy (Crozier et al., 2009). Although statistically significant, the identified difference in maternal age (pregnant: 33 ± 4 and non-pregnant: 30 ± 7) was not considered biologically relevant since the difference of the mean values was only 3 years.

Overall microbial profiles could not be distinguished from each other based on pregnancy status alone. However, a more detailed analysis revealed several differences between the microbiomes of healthy pregnant women and those of healthy non-pregnant women. The BV-associated CST IVA (dominated by *G. vaginalis* subgroup B and *Atopobium*) was not detected in the pregnant cohort, whereas *L. gasseri* dominated CST II, which was not detected among the 310 non-pregnant women in our previous study (Albert et al., 2015), was detected in 12 women in the pregnant cohort. Pregnant women in *Lactobacillus*-dominated CST had higher relative abundance of *Lactobacillus* spp. when compared to non-pregnant women. Vaginal microbiomes of pregnant women had lower richness and diversity and a correspondingly lower prevalence of Mollicutes and *Ureaplasma* when compared to non-pregnant women.

Microbial profiles from the pregnant women clustered in six different groups, mostly *Lactobacillus*-dominated CST (CST I, CST II, CST III and CST V), originally defined by Ravel & Gajer (Brotman et al., 2014b). Non-*Lactobacillus*-dominated (CST IV) profiles are described in the literature as either very heterogeneous or dominated with BV-associated bacteria (Romero et al., 2014b; Albert et al., 2015). None of the pregnant women, including 61 with intermediate or high Nugent scores, were identified as belonging to CST IVA, which is characterized by dominance of *G. vaginalis* subgroup B and *Atopobium* (Albert et al., 2015). This distribution is notably different from the non-pregnant cohort, where 24.6% of women with intermediate or high Nugent scores were assigned to CST IVA (Albert et al., 2015). Results of other studies have suggested that CST dominated by BV-associated microorganisms are less frequently detected in pregnancy (Romero et al., 2014c; MacIntyre et al., 2015). While our study design does not allow us to address the issue of overall prevalence of BV-associated CST in pregnancy, our results suggest that the distribution of these CST among pregnant women may be different than in non-pregnant women. This suggests a role CST IVA could be playing in early pregnancy loss.

OTUs that were weakly similar to *Streptococcus* and *Weissella* species were detected in most samples, but they represented only 0.33% and 0.22% of all reads, respectively. These OTUs were previously described as highly prevalent in the vaginal microbiome of healthy non-pregnant women (Albert et al., 2015). They have low sequence identity to any reference sequences in the cpnDB_nr database (OTU 0026: *S. devriesei* 83% and OTU 0021: *W. viridescens* 58.8%). However, this subset of the cpnDB database contains only selected representative sequences of named species. A broader search shows that these OTU are more similar to metagenomic

sequences derived from the fecal microbiome (OTU 0021 is 97% identical to Genbank accession GQ178631) or oral microbiome (OTU 0026 is 85% identical to KJ406686) that represent uncharacterized Firmicutes; reminders of the common but still uncharacterized constituents of the human microbiome.

Our findings of greater *Lactobacillus* abundance and lower richness and diversity in the vaginal microbiomes of pregnant women relative to non-pregnant women are consistent with previous studies in the literature. Aagaard et al. (Aagaard et al., 2012) analyzed the microbiomes of 24 healthy, pregnant American women sampled at three different locations within the vagina. Vaginal site did not drive the structure of the microbial community, but the authors found that overall microbiomes of pregnant women were less diverse and less rich when compared to non-pregnant women. Similarly, Walther-António et al. (Walther-António et al., 2014) described the microbiomes of 12 White pregnant American women based on longitudinal sampling and observed reduced microbiome diversity and higher *Lactobacillus* spp. relative abundance during pregnancy. Romero et al. (Romero et al., 2014c) also reported that *Lactobacillus* spp. abundance was significantly higher in pregnant women in comparison to non-pregnant and increased as a function of gestational age. They also described higher stability of the microbiome during pregnancy when compared to reproductive age non-pregnant women. In another longitudinal study, MacIntyre et al. (MacIntyre et al., 2015) analyzed the microbiomes of 42 British women during pregnancy and the post-partum period. *L. jensenii*-dominated profiles were more common among these women than Northern American women. The authors also observed that post-partum microbiomes become less *Lactobacillus* spp. dominant and more rich and diverse (i.e. more similar to the microbiomes of non-pregnant women) regardless of ethnicity, providing strong support for the idea that pregnancy has a transient effect on the vaginal microbial community. Importantly, the conclusions of these studies and our current study are consistent despite differences in the cohort studied (Canadian, American or European cohorts of varying mixtures of ethnicity), universal target amplified (cpn60 universal target or 16S rRNA gene) or sequencing platforms used (454/Roche pyrosequencing or Illumina MiSeq).

The explanations for these pregnancy-associated changes are not well established, but a relationship between sex steroid hormone levels and the composition of the vaginal microbiome has been previously reported (Brotman et al., 2014b; Devillard et al., 2004; Farage et al., 2010). Increased levels of estrogen during pregnancy lead to increased thickness of the vaginal mucosa

and increased deposition of glycogen (BMJ, 1943; Cruickshank and Sharman, 1934). Glycogen is the main carbohydrate utilized by *Lactobacillus* spp. for the production of lactic acid, which contributes to the protective effect of a low vaginal pH (Mirmonsef et al., 2014; Prince et al., 2014; O'Hanlon et al., 2013; Cruickshank, 1934). This may contribute to the greater dominance of *Lactobacillus* in pregnancy and, consequently, the lower richness and diversity in this cohort.

A novel finding in our study was the lower prevalence of Mollicutes and *Ureaplasma* in pregnancy detected by family and genus-specific PCR. Mollicutes have been associated with preterm birth, preterm premature rupture of membranes and low birth weight (Perni et al., 2004; Foxman et al., 2014; Kwak et al., 2014). The lower prevalence of Mollicutes and *Ureaplasma* is consistent with the overall lower species richness and diversity in the vaginal microbiomes of pregnant relative to non-pregnant women. We also found that pregnant women had higher bacterial load than non-pregnant women as estimated by quantitative PCR targeting the 16S rRNA gene. Hormone induced production of glycogen may offer a nutritionally richer environment for bacterial growth in the vagina during pregnancy. Additionally, it is known that pregnancy alters the amount and consistency of the mucus, which becomes more abundant and thicker (Chrétien, 1978). Thus, it is possible that swabs sampled from the pregnant women carried more material when compared to non-pregnant women. We are unable to resolve this question since the swabs were not weighed before DNA extraction steps. In addition to pregnancy associated physiological differences and mucus consistency, other technical factors such as storage conditions or inherent differences in the study populations cannot be ruled out.

One limitation of this study was the assessment of pregnancy outcomes and microbial profiles since there were very few poor outcomes in this cohort, as it would be expected for a low risk group. This study, however, does provide crucial baseline information for future studies in pregnant women. In addition, we detected significant interaction between parity and CST. Considering the large number of variables in the metadata, analysis of these associations should be interpreted with caution. We can speculate that the prevalence of the *L. crispatus*-dominated CST I among nulliparous women might be associated with more cautious or health conscious behaviour among women in their first gestation. Pregnancy-induced physiologic alterations that can persist after delivery have been previously reported (Clifton et al., 2012). In addition to physiologic changes, a disturbed vaginal microbiome that persisted for up to a year post-partum has also recently been described (DiGiulio et al., 2015). Those persistent changes might explain

the association between CST and parity we observed, with post-partum microbiome changes affecting the current status of primiparous and multiparous women. Further studies are needed to investigate these relationships in more detail.

In conclusion, we have identified several differences in the composition of the vaginal microbial communities of pregnant women living in Canada relative to non-pregnant women: larger total bacterial community, lower richness and diversity, higher *Lactobacillus* abundance and lower Mollicutes/*Ureaplasma* prevalence. These findings give us a better understanding of the vaginal microbiome in pregnancy, which is a critical step toward being able to exploit the diagnostic potential of the microbiome for the prediction of adverse pregnancy outcomes as well as to explore alternative therapeutic procedures through microbiological intervention. Establishing an understanding of the normal microbiome in low risk pregnant women is a vital baseline for comparison to the microbiome of women who have adverse perinatal outcomes such as preterm birth.

Acknowledgements

The authors are grateful to the women who participated in the study and to the research staff at Mount Sinai Hospital. Financial support was provided by a joint Canadian Institutes of Health Research (CIHR) Emerging Team Grant and a Genome British Columbia (GBC) grant awarded to DMM, AB and JEH (grant reference #108030) as well as a CIHR grant MOP-82799 to AB. ACF was supported by a University of Saskatchewan graduate scholarship.

VOGUE Research Group: Sean Hemmingsen, Gregor Reid, Tim Dumonceaux, Gregory Gloor, Matthew Links, Kieran O'Doherty, Patrick Tang, Julianne van Schalkwyk and Mark Yudin.

Table 2.4S. Supplementary material (.xls). Summary of OTU analyzed in this study.
Best database match, percentage of identity, cpnDB name, taxonomic lineage and abundance in each library are shown.

Available at: https://www.nature.com/articles/s41598-017-07790-9#Sec19

CHAPTER 3. Increased richness and diversity of the vaginal microbiota and spontaneous preterm birth

Citation

Freitas AC, Bocking A, Hill JE, Money DM & The VOGUE Research Group. Increased richness and diversity of the vaginal microbiota and spontaneous preterm birth. (Submitted, October 19, 2017).

Author contributions

Conceived the study: DM, AB, JEH and the other members of the VOGUE Research Group. Oversaw and contributed to data collection and analysis, and participated in manuscript review: DM, AB, and JEH. Performed the experiments, analyzed the data and wrote the paper ACF.

Chapter transition

Preterm birth affects millions of babies every year and preterm birth complications are the leading cause of neonatal mortality. Preterm birth is a multifactorial condition and an intra-uterine infection ascending from the vagina is thought to be one of the potential contributors to preterm birth. In the previous chapter, I established a baseline description of the vaginal microbiome in healthy pregnancies. Having the information of how the microbiome in a "healthy state" in pregnancy looks like provided me the opportunity to investigate the effect of the microbiota in preventing or leading to preterm birth. The goal was to identify traits in the microbiome of women who had spontaneous preterm birth that could be potentially used as diagnostic targets. The results of this study showed that the microbiomes of women who experienced preterm birth have higher richness and diversity, and Mollicutes prevalence than the microbiomes of women with term deliveries. At the overall level, I did not identify a core microbiome or specific traits that could predict outcome, i.e., preterm birth. The differences in the microbial profiles of women with term and preterm deliveries, however, suggest that the microbiome indeed is somehow associated with reproductive outcomes.

3.1. Abstract

The bacterial community present in the female lower genital tract plays an important role in maternal and neonatal health. Imbalances in this microbiota have been associated with negative reproductive outcomes, such as spontaneous preterm birth (sPTB), but the mechanisms underlying the association between a disturbed microbiota and sPTB remain poorly understood. An intrauterine infection ascending from the vagina is thought to be an important contributor to the onset of preterm labour. Our objective was to characterize the vaginal microbiota of pregnant women who had sPTB (n=46) and compare to those of pregnant women who delivered at term (n=170). Vaginal swabs were collected from women at 11-16 weeks of gestational age. Microbiota profiles were created by PCR amplification and pyrosequencing of the cpn60 universal target region. Profiles clustered into seven Community State Types: I (*Lactobacillus crispatus* dominated), II (*Lactobacillus gasseri* dominated), III (*Lactobacillus iners* dominated), IVA (*Gardnerella vaginalis* subgroup B or mix of species), IVC (*G. vaginalis* subgroup A dominated), IVD (*G. vaginalis* subgroup C dominated) and V (*Lactobacillus jensenii* dominated). The microbiota of women who experienced preterm birth (<37 weeks gestation) had higher richness and diversity, and higher Mollicutes prevalence when compared to those of women who delivered at term. The two groups did not cluster according to CST, likely because CST assignment is driven in most cases by the dominance of one particular species, overwhelming the contributions of more rare taxa. In conclusion, we did not identify a specific microbial community structure that predicts sPTB, but differences in microbiota richness, diversity and Mollicutes prevalence were observed between groups. Although a causal relationship remains to be determined, our results confirm previous reports of an association between Mollicutes and sPTB, and further suggest that a more diverse microbiome may be important in the pathogenesis of some cases.

Keywords: microbiome, vagina, *Lactobacillus*, CST, diversity, richness, Mollicutes, preterm birth, pregnancy, infection.

3.2. Introduction

Preterm birth is defined as delivery before 37 completed weeks of gestational age (WHO *et al.*, 2012) and can be further sub-categorized in: extremely preterm ($\leq 27^{+6}$ weeks^{+days}), very preterm (28 to 31^{+6}), and late preterm (32 to 36^{+6}) (Quinn et al., 2016). Preterm birth comprises

11% of all livebirths worldwide and its complications are estimated to cause 35% of world's neonatal deaths, which represents 3.1 million deaths annually (Blencowe et al., 2012). Children who are born prematurely also have higher rates of cardiovascular disorders, respiratory distress syndrome, neurodevelopmental disabilities and learning difficulties compared with children born at term (Behrman and Butler, 2007).

Preterm birth is a complex multi-factorial condition with several known risk factors, such as low and high maternal ages (Fraser et al., 1995; Jacobsson et al., 2004; Schempf et al., 2007), low BMI (Han et al., 2011), black ethnicity (Kistka et al., 2007), tobacco use, heavy alcohol intake, illicit drug use (Behrman and Butler, 2007), close temporal proximity to a previous delivery (Conde-Agudelo et al., 2006), and multiple gestation (Stock and Norman, 2010). Although studied extensively, some preterm cases remain unexplained for women with no known risk factors. Intrauterine infection with organisms ascending from the vagina has been hypothesized as an important contributor to preterm birth since many organisms isolated from the amniotic fluid/membranes of women who experienced preterm are also found in the lower genital tract of pregnant women (Gardella et al., 2004; Krohn et al., 1995; Hillier et al., 1988; Romero et al., 1989). A large number of studies support this hypothesis based on the strong association between intra-amniotic bacterial infection and preterm birth (Gardella et al., 2004; Hillier et al., 1988; Romero et al., 1989; Watts et al., 1992; Yoon et al., 2001; DiGiulio et al., 2008; Hillier et al., 1993; Jacobsson et al., 2005).

The microbiological diagnosis of a "normal" or disturbed vaginal microbiota has historically been based on the Nugent score, the current gold standard diagnostic method that relies on Gram stain of vaginal smears (Nugent et al., 1991). The "normal" vaginal microbiota in non-pregnant reproductive aged women is understood to be dominated by *Lactobacillus* species, while an abnormal microbiota (defined as Bacterial vaginosis) is characterized by low abundance of lactobacilli and an overgrowth of anaerobic bacteria, such as *Gardnerella vaginalis*, *Prevotella* spp., *Bacteroides* spp., *Mobiluncus* spp., and *Mycoplasma hominis* (Hill, 1993). In low risk pregnant women, it has been shown that the vaginal microbiota has reduced richness and diversity, and increased abundance of lactobacilli compared to those of non-pregnant women (Romero et al., 2014c; Freitas et al., 2017; Aagaard et al., 2012; Walther-António et al., 2014; MacIntyre et al., 2015). An abnormal microbiota has been previously associated with preterm birth (Leitich et al., 2003), but only a few in depth culture independent studies of the vaginal microbiota of women

who had preterm birth have been published, with inconsistent conclusions (Romero et al., 2014b; Hyman et al., 2014; DiGiulio et al., 2015).

The objective of this study was to assess whether there are differences in the vaginal microbiota composition of women who had spontaneous preterm birth (sPTB) and term delivery early in gestation that could be further investigated as diagnostic indicators of preterm birth risk. Microbiome profiling was based on sequencing of the cpn60 universal target, which provides higher resolution than 16S rRNA variable regions (Links et al., 2012) and allows the resolution of *Gardnerella vaginalis* subgroups, a hallmark bacterium in the disturbed microbiota (Paramel Jayaprakash et al., 2012).

3.3. Methods

Study population and sampling

This study received ethical approval from the University of British Columbia Children's and Women's Research Ethics Board (Approval Number H14-01954), and Mount Sinai Hospital Research Ethics Board (Approval Number 15-0184-E). All participants provided written consent at enrolment.

This retrospective cohort study analyzed the vaginal microbiota of women who experienced spontaneous preterm birth (sPTB) and compared the resulting microbial profiles to those of pregnant women who delivered at term. The bacterial profiles of pregnant Canadian women at low-risk of sPTB who had term deliveries (n=170) were previously generated by our research group (Freitas et al., 2017). The vaginal microbial profiles of Canadian women who had preterm birth originated from samples of this previous study (n=7) (Freitas et al., 2017), and from the Ontario Birth Study (n=39), resulting in 46 samples. The Ontario Birth Study is an open longitudinal pregnancy cohort at Mount Sinai Hospital, Toronto, Canada. It is a platform for studies of both pregnancy complications as well as Developmental Origins of Health and Disease related research. All biospecimens, including maternal vaginal swabs and maternal and infant blood, are collected concurrently with routine clinical specimens to reduce the burden on study participants. Detailed demographic and lifestyle characteristics are obtained from women during pregnancy and postpartum and clinical information is extracted from the health records. For the

purposes of this report, self-administered vaginal swabs were taken at 16 weeks gestation and placed in dry tubes prior to being placed in -80˚C for storage in the Lunenfeld Tanenbaum Research Institute Biospecimen Storage and processing Laboratory. Specimens from all cohorts were processed similarly in terms of sample collection, storage, DNA extraction, library preparation and sequencing.

Clinical and behavioural questionnaire data (pregnancy history, family and personal medical history, psychosocial health, demographic factors and other lifestyle and environmental exposures) were transferred to the Research Electronic Data Capture (REDCap) database protected by a secure server (Harris et al., 2009). For the PTB group, eligible participants for this study were women who had undergone preterm delivery at greater than 20 weeks but less than 37 weeks gestational age, where onset of labour occurred spontaneously or in association with cervical incompetence or preterm premature rupture of membranes (PPROM). Vaginal swabs collected from pregnant women (both PTB and term groups) at 11-16 weeks of gestational age were used for bacterial genomic analysis

Total nucleic acid was extracted from swabs using the MagMAX™ Total Nucleic Acid Isolation Kit (Life Technologies, Burlington, ON, Canada) as per manufacturer's instructions. Kit reagents are aliquoted to eliminate repeated accessing of open reagents, and samples are processed in small batches using filter-tips to prevent cross-contamination. Pipettes and other lab surfaces are regularly treated with DNA surface decontaminant (DNA Away, Thermo Fisher Scientific, Waltham, MA).

Total Bacterial DNA (qPCR) and Detection of Mollicutes (PCR)

Quantitative PCR (qPCR): Total bacterial DNA quantity in each sample was estimated using a SYBR Green assay based on amplification of the V3 region of the 16S rRNA gene. Primer sequences were as follows: SRV3-1 (5'-CGGYCCAGACTCCTAC-3'), SRV3-2 (5'-TTACCGCGGCTGCTGGCAC-3') (Lee et al., 1996). Reactions run on a MyiQ thermocycler using the following cycling parameters: 95°C for 3 min, followed by 30 cycles of 95°C for 15 sec., 62°C for 15 sec., 72°C for 15 sec., with a final extension at 72°C for 5 minutes (Chaban et al., 2013).

Conventional PCR: Some Mollicutes (*Mycoplasma* and *Ureaplasma*) species lack a cpn60 gene (Hill et al., 2004). Thus, we performed a family-specific semi-nested PCR targeting the 16S rRNA gene to detect Mollicutes (van Kuppeveld et al., 1992), and a PCR targeting the multiple banded antigen gene to detect *Ureaplasma* spp.. PCR products from *U. parvum* and *U. urealyticum* can be differentiated by size (Watson et al., 1990).

cpn60 Universal Target (UT) PCR and Pyrosequencing

Universal primer PCR targeting the 549-567 bp cpn60 UT region was performed using a mixture of cpn60 primers consisting of a 1:3 molar ratio of primers H279/H280:H1612/H1613, as described previously (Chaban et al., 2014; Hill et al., 2006b; Schellenberg et al., 2011). To allow multiplexing of samples in a single sequencing run, primers were modified at the 5' end with one of 24 unique decamer multiplexing identification (MID) sequences, as per the manufacturer's recommendations (Roche, Brandford, CT, USA). Amplicons were pooled in equimolar amounts for sequencing on the Roche GS Junior sequencing platform. The sequencing libraries were prepared using the GS DNA library preparation kit and emulsion PCR (emPCR) was performed with a GS emPCR kit (Roche Diagnostics, Laval, Canada).

Samples were handled in small batches to avoid cross-contamination and experimental controls were included at several steps in the study. Regular monitoring DNA extraction controls in our lab by universal PCR confirms that these procedures are sufficient to eliminate detectable template contamination of study samples. A no template control was also included in each set of PCR reaction as negative controls. Experimental controls were not sequenced as they did not yield any amplification.

Analysis of Operational Taxonomic Units (OTU)

Raw sequence data was processed by using the default on-rig procedures from 454/Roche. Filter-passing reads were used in the subsequent analyses for each of the pyrosequencing libraries. MID-partitioned sequences were mapped using Bowtie 2 (http://bowtie-bio.sourceforge.net/bowtie2/) on to a manually curated reference set of 1,561 OTU sequences

representing the human vaginal microbiota. Bowtie 2 was run using the default end-to-end alignment mode.

The OTU reference set was generated originally by de novo assembly of cpn60 sequence reads from 546 vaginal microbiomes using the microbial Profiling Using Metagenomic Assembly pipeline (mPUMA, http://mpuma.sourceforge.net) (Links et al., 2013) with Trinity as the assembly tool (Grabherr et al., 2011). OTU were labeled according to their nearest reference sequence determined by watered-Blast comparison (Schellenberg et al., 2009) to the cpn60 reference database, cpnDB_nr (downloaded from http://www.cpndb.ca (Hill et al., 2004)). This reference assembly approach allows us to compare the microbial profiles from various cohorts under investigation, including the 46 pregnant women who had sPTB described in this study.

The result of mapping is an OTU frequency table (Supplementary Table 3.4S) that was used for microbiome data analysis. Some analyses were also performed at species level, i.e., combined OTU that have the same nearest neighbour.

Statistical Analysis

Comparisons of socio-demographic characteristics of cohorts and participants were based on analysis of variance (ANOVA), t-test and Chi-square, performed in IBM SPSS (Statistical Package for the Social Sciences, version 21) at 5% level of significance. For analysis of associations between socio-demographic characteristics and microbiota profiles (CST), a false discovery rate (FDR) correction for multiple comparisons was applied (Benjamini and Hochberg, 1995).

Alpha (Shannon diversity and Chao1 estimated species richness) and beta diversity (jackknifed Bray-Curtis dissimilarity matrices) were calculated as the mean of 100 subsamplings of 1000 reads (or all reads available when less than 1000) in QIIME (Quantitative Insights Into Microbial Ecology) (Caporaso et al., 2010). Plots of alpha diversity measures against bootstrap sample number were generated in R and visually inspected to ensure that an adequate sampling depth for each sample was achieved.

For community state type (CST) analysis, a Jensen-Shannon distance matrix was calculated in R with a custom distance function that calculates the square root of the Jensen-Shannon

divergence (Endres and Schindelin, 2003). This distance matrix was used for hierarchical clustering using the 'hclust' function in R, with Ward linkage.

The function aldex.clr from the ALDEx2 package in R was used to compare the differential relative abundance of individual taxa in term and preterm groups (Fernandes et al., 2013). Significant differences were determined based on the false discovery rate (FDR), which is the result of a Benjamini-Hochberg corrected p-value from a Welch's t-test calculated within ALDEx2.

3.4. Results

Description of the study population and pregnancy outcomes

Socio-demographic characteristics of women who had spontaneous preterm birth (n=46) and women who had term deliveries (n=170) are summarized in Table 3.1. There were no significant differences in maternal age, BMI, ethnicity, smoking status, consumption of alcohol or use of probiotics between term and preterm groups (all $p>0.05$). Average maternal age was 33 for participants in both cohorts. Average body mass index (BMI) was 22.9 and 24.2 for women in the term and preterm groups, respectively. Most women in both cohorts identified themselves as white ethnicity, followed by East Asian and South/Southeast Asian (Table 3.1). Consumption of tobacco (term 2.3%; preterm 0%), alcohol (term 5.9%; preterm 4.3%) or probiotic supplements (term 4.1%; preterm 6.5%) was low among women in both groups (chi-square, all $p>0.05$).

Most women in the preterm group had a Bachelor/graduate degree (29/46) and an average house income higher than CAD 100000 per year (25/46). A minority of women who had preterm birth (5/46) reported consumption of substances without prescription prior pregnancy, of which 3/46 women consumed marijuana/hashish, 1/46 woman consumed tranquilizers/nerve pills and 1/46 woman consumed cocaine/crack. Approximately 74% of the participants in the preterm group reported a pre-existing condition. A total of 12/46 women had some condition related to mental health, such depression or anxiety. Twenty four percent of women (11/46) had a genitourinary condition and 17% (8/46) had neurological condition, such as migraines.

Table 3.1. Socio-demographic and microbiological characteristics of subjects.

Characteristics	Term pregnancies (n=170)	Preterm pregnancies (n=46)	p-value
Age [Mean ± SD (Range)] [1*]	33.6 ± 4.2 (21 - 45)	33.65 ± 4.1 (25 - 45)	0.948
21-25	5 (2.9%)	1 (2.1%)	
26-35	114 (67.1%)	32 (69.5%)	
36-45	51 (30.0%)	13 (28.2%)	
BMI [Mean ± SD (Range)] [1]	22.9 ± 3.8 (17 - 40)	24.2 ± 5.6 (19 - 43)	0.125
Underweight (<18.50)	7 (4.1%)	0 (0%)	
Normal weight (18.51-24.9)	131 (77.0%)	33 (73.3%)	
Overweight (25.0-29.9)	25 (14.7%)	8 (17.7%)	
Obese (>30)	7 (4.1%)	4 (9.0%)	
MD	0 (0%)	1 (2.2%)	
Ethnicity [2]			0.261
White	108 (63.5%)	22 (47.8%)	
East Asian	26 (15.3%)	6 (13.0%)	
South / Southeast Asian	15 (8.8%)	4 (8.7%)	
Latin America / Hispanic	8 (4.7%)	3 (6.5%)	
Black	3 (1.8%)	2 (4.4%)	
Other/Mixed ethnicity	10 (5.9%)	6 (13.0%)	
MD	0 (0%)	3 (6.5%)	
Community state type (CST) [2]			0.361
I	56 (32.9%)	17 (37%)	
II	9 (5.3%)	5 (10.9%)	
III	28 (16.5%)	8 (17.4%)	
IVA	31 (18.2%)	6 (13%)	
IVC	19 (11.2%)	2 (4.3%)	
IVD	11 (6.5%)	1 (2.2%)	
V	16 (9.4%)	7 (15.2%)	
Estimated bacterial load (log copies of 16S rRNA gene) / swab [Mean ± SD (Range)][1]	7.78 ± 0.93 (4.89 - 10.67)	8.07 ± 0.71 (6.32 - 10.33)	0.049
Presence of Mollicutes [2]	68 (40%)	28 (60.8%)	0.012
Presence of *Ureaplasma* [2]	40 (23.4%)	14 (30.4%)	0.337
U. parvum	37 (21.7%)	14 (30.4%)	
U. urealyticum	3 (1.7%)	0 (0%)	
Shannon diversity [Mean ± SD (Range)] [1]	1.28 ± 0.86 (0.13 - 4.52)	1.81 ± 1.13 (0.34 - 5.16)	0.004
Chao1 richness [Mean ± SD (Range)] [1]	36.22 ± 14.80 (14.39 - 115.74)	46.38 ± 24.19 (20.20 - 126.01)	0.009

*[1]t-test; [2]Chi-square. MD = missing data.

Characteristics regarding pregnancy and neonatal outcomes are described in Table 3.2. Pregnancy outcome information was not available for one woman in the preterm group as she was lost to follow-up. There were no significant differences in gestational age at enrolment, mode of conception or fetal sex between groups (all $p>0.05$). Average gestational age at delivery was 39^{+3} weeks for the women who delivered at term and 34^{+2} weeks for women who had preterm birth, most of which were considered late preterm, i.e., delivery between 32 to 36^{+6} weeks of gestational age. Women in the preterm group were more likely to have experienced preterm birth or miscarriage in their previous pregnancy (Chi-square, $p<0.001$). They also had higher percentage of caesarean sections than women who delivered at term. Parity also differed between groups; women who had preterm birth were more likely to be nulliparous. There was a significant difference between term and preterm groups regarding birth weight and number of infants admitted to level 3 neonatal intensive care unit (NICU) (Table 3.2). Apgar score at 1 (term 8.75 ± 0.6; preterm 8.38 ± 1.1) and 5 minutes (term 8.97 ± 0.17; preterm 8.76 ± 0.7) between groups also differed (t-test, all $p<0.001$). One preterm infant (1/46) died shortly after birth (20 weeks of gestational age).

Among women who delivered preterm, 56% (26/46) had premature rupture of membranes (PPROM), 10.8% (5/46) had gestational diabetes and 4.3% (2/46) had anemia unresponsive to therapy. Twenty-four percent of women (11/46) presented one of the following conditions: maternal elevated liver enzymes, short cervix and incompetent cervix; fetal ascites, fetal distress, large fetus for gestational age; placental findings of marginal cord insertion, two vessel umbilical cord, placenta previa, low lying placenta.

Sequencing results and OTU analysis

Raw sequence data files for the samples described in this study were deposited to the NCBI Sequence Read Archive (Accession SRP073152, BioProject PRJNA317763; BioProject PRJNA403856). Total dataset contained 1,635,072 cpn60 reads; median and average read count per sample were 4936 and 7569 (range 402 – 37378), respectively. Average read length was 424 bp. Results of Bowtie2 mapping showed that these reads corresponded to 728 OTUs from the reference assembly (Supplementary Table 3.4S).

Table 3.2. Gestation characteristics, pregnancy and neonatal outcomes.

Characteristics	Term pregnancies (n=170)	Preterm pregnancies (n=46)	p-value
Gestational age in weeks^{+day}			
At enrolment [Mean ± SD (Range)] [1]	$13^{+2} \pm 1^{+1}$ (11^{+1} - 16^{+6})	$13^{+3} \pm 1^{+0}$ (11^{+6} - 16^{+0})	0.641
At delivery [Mean ± SD (Range)] [1]	$39^{+3} \pm 0^{+6}$ (39^{+3} - 41^{+2})	$34^{+2} \pm 2^{+6}$ (20^{+0} - 36^{+6})	<0.0001
Late preterm (32 to 36^{+6})	NA	39 (84.8%)	
Very preterm (28 to 31^{+6})	NA	5 (10.8%)	
Extremely early ($\leq 27^{+6}$)	NA	1 (2.1%)	
Previous pregnancy history (excludes women in first pregnancy) [2]	(n=125)	(n=24)	<0.0001
Livebirth, term	89 (71.2%)	9 (37.5%)	
Livebirth, preterm	0 (0%)	3 (12.5%)	<0.0001
Spontaneous abortion	20 (16%)	8 (33.3%)	0.046
Pregnancy termination	16 (12.8%)	3 (12.5%)	0.968
Ectopic pregnancy	0 (0%)	1 (4.1%)	
Mode of delivery [2]			0.042
Vaginal delivery	128 (75.3%)	27 (58.7%)	
C-section	42 (24.7%)	18 (39.1%)	
Parity [2]			0.005
0	55 (32.3%)	28 (60.8%)	
1	94 (55.3%)	12 (26%)	
2-4	21 (12.3%)	6 (13%)	
Assisted conception [2]	17 (10%)	6 (13%)	0.591
Fetal sex (% male / % female) [2]	(48.8%) / (51.1%)	(41.3%) / (56.5%)	0.430
Birth weight (g) [Mean ± SD (Range)] [1]	3398 ± 459 (1970 - 5200)	2550 ± 559 (1300 - 3595)	<0.0001
Infant in NICU (n, %) [2]	1 (0.6%)	24 (52.2%)	<0.0001

*[1]t-test; [2]Chi-square.

Microbiota profiles

Microbiota profiles were created by PCR amplification and pyrosequencing of the cpn60 universal target region. Hierarchical clustering of vaginal microbiota profiles resulted in seven Community State Types (CST): I (*Lactobacillus crispatus* dominated), II (*Lactobacillus gasseri* dominated), III (*Lactobacillus iners* dominated), IVA (*Gardnerella vaginalis* subgroup B or mix of different species), IVC (*G. vaginalis* subgroup A dominated), IVD (*G. vaginalis* subgroup C dominated) and V (*Lactobacillus jensenii* dominated) (Figure 3.1). Each CST is defined by the dominance of one species of *Lactobacillus* (I, II, III, IV), *Gardnerella vaginalis* (IVC, IVD) or a mixture of bacteria species (IVA), as previously described (Ravel et al., 2011; Albert et al., 2015).

Overall microbiota profiles did not cluster together based on gestational age at delivery (Figure 3.1 and 3.2). Most microbial profiles from the preterm group (80.5%) were assigned to *Latobacillus*-dominated CST: CST I (37% of profiles), CST III (17.4%), CST V (15.2%), and CST II (10.9%). The remaining profiles (19.5%) were assigned to CST IVA, IVC or IVD (Table 3.1). The CST IVA was the most heterogeneous group, represented by the dominance of *Lactobacillus delbrueckii, Bifidobacterium dentium, Bifidobacterium infantis, Atopobium vaginae, Bifidobacterium breve* or a mixture of different bacteria species. The CST IVC was dominated by *G. vaginalis* subgroup A and *Megasphaera* spp., and CST IVD was dominated by *G. vaginalis* subgroup C (Figure 3.1).

Ecological analysis and total bacterial load

Assessment of alpha diversity revealed that microbiomes of women who delivered preterm were richer (Chao1 richness 46.3 ± 24.1) and more diverse (Shannon diversity index 1.8 ± 1.1) when compared to those of women in the term group (36.2 ± 14.8; 1.2 ± 0.8) (t-test, $p<0.01$) (Table 3.1). Total bacterial load was estimated based on qPCR targeting the 16S rRNA gene and it was expressed as log 16S rRNA gene copy number per swab. Higher bacterial loads were detected in samples from the preterm group (7.7 ± 0.9) compared to term group (8.0 ± 0.7) (t-test, $p=0.049$) (Table 3.1).

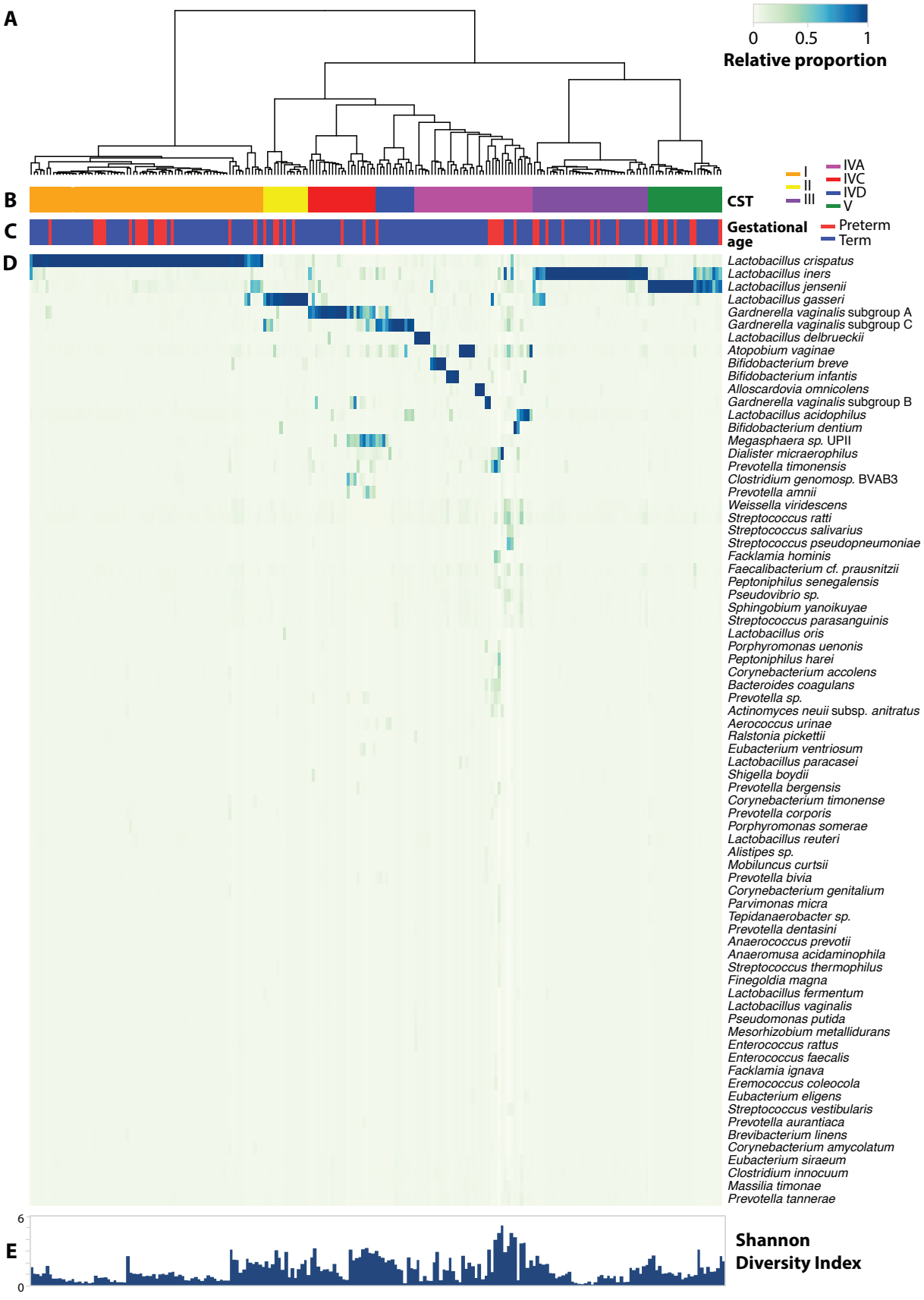
A
0
0.5
1
Relative proportion
B
CST
I
II
III
IVA
IVC
IVD
V
C
Gestational age
Preterm
Term
D
Lactobacillus crispatus
Lactobacillus iners
Lactobacillus jensenii
Lactobacillus gasseri
Gardnerella vaginalis subgroup A
Gardnerella vaginalis subgroup C
Lactobacillus delbrueckii
Atopobium vaginae
Bifidobacterium breve
Bifidobacterium infantis
Alloscardovia omnicolens
Gardnerella vaginalis subgroup B
Lactobacillus acidophilus
Bifidobacterium dentium
Megasphaera sp. UPII
Dialister micraerophilus
Prevotella timonensis
Clostridium genomosp. BVAB3
Prevotella amnii
Weissella viridescens
Streptococcus ratti
Streptococcus salivarius
Streptococcus pseudopneumoniae
Facklamia hominis
Faecalibacterium cf. prausnitzii
Peptoniphilus senegalensis
Pseudovibrio sp.
Sphingobium yanoikuyae
Streptococcus parasanguinis
Lactobacillus oris
Porphyromonas uenonis
Peptoniphilus harei
Corynebacterium accolens
Bacteroides coagulans
Prevotella sp.
Actinomyces neuii subsp. anitratus
Aerococcus urinae
Ralstonia pickettii
Eubacterium ventriosum
Lactobacillus paracasei
Shigella boydii
Prevotella bergensis
Corynebacterium timonense
Prevotella corporis
Porphyromonas somerae
Lactobacillus reuteri
Alistipes sp.
Mobiluncus curtsii
Prevotella bivia
Corynebacterium genitalium
Parvimonas micra
Tepidanaerobacter sp.
Prevotella dentasini
Anaerococcus prevotii
Anaeromusa acidaminophila
Streptococcus thermophilus
Finegoldia magna
Lactobacillus fermentum
Lactobacillus vaginalis
Pseudomonas putida
Mesorhizobium metallidurans
Enterococcus rattus
Enterococcus faecalis
Facklamia ignava
Eremococcus coleocola
Eubacterium eligens
Streptococcus vestibularis
Prevotella aurantiaca
Brevibacterium linens
Corynebacterium amycolatum
Eubacterium siraeum
Clostridium innocuum
Massilia timonae
Prevotella tannerae
E
6
0
Shannon Diversity Index

Figure 3.1. Vaginal microbiota profiles of women who had sPTB and term deliveries.

(A) Hierarchical clustering of Jensen-Shannon distance matrices with Ward linkage on the relative proportions of reads for each OTU within individual vaginal samples. (B) Community state type (CST). (C) Gestational age at delivery. (D) Heatmap of relative abundances of bacterial species within each vaginal microbiota. Each column represents a woman's vaginal microbiota profile, and each row represents a bacteria species. Only species that are at least 1% abundant in at least one sample are shown. (E) Shannon diversity indices calculated for each sample.

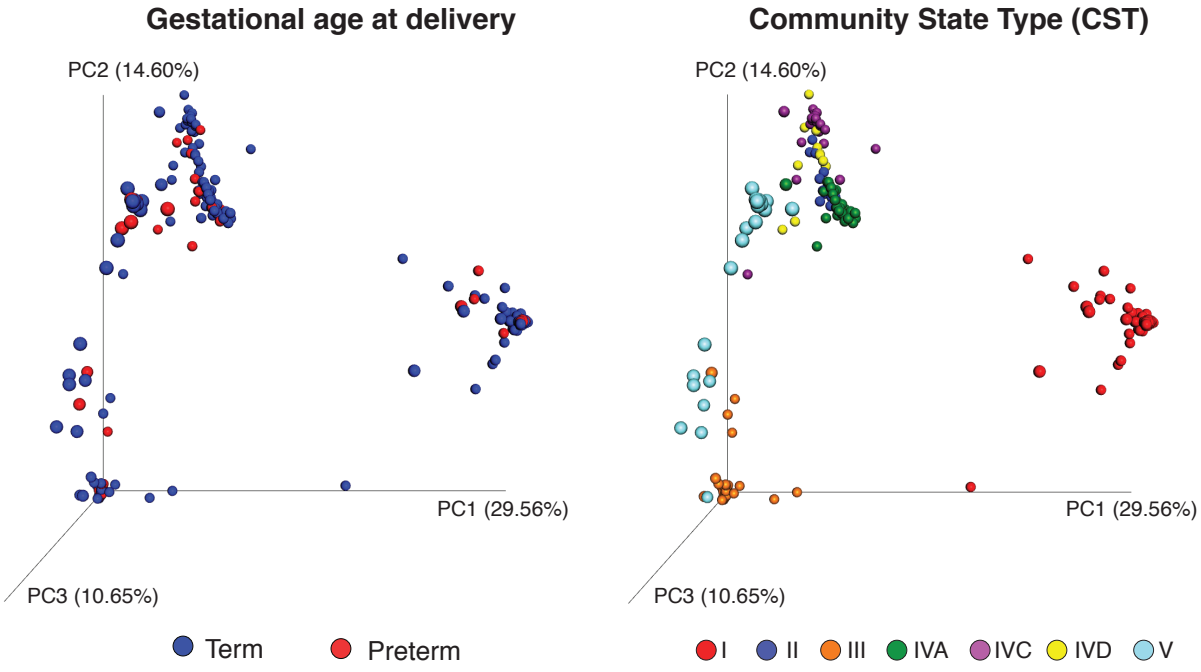

Figure 3.2. Vaginal microbiota profiles coloured by gestational age at delivery or CST.

Jackknifed principal coordinates analysis (PCoA) of Bray-Curtis distance matrices of microbial profiles from all participants in the study.

Bacterial species relative abundance and prevalence

To investigate whether there was an association between individual taxa and sPTB, the abundance and prevalence of each species was evaluated. The ALDEx2 analysis assessed the relative abundance of each taxa (at the OTU and species level) in term and preterm groups. Eight OTU/species were more abundant in the term group in comparison with preterm, all of which were considered rare members of the bacterial community (Figure 3.3). *L. acidophilus* represented 1% of the total reads in the dataset and had a low relative abundance average of 1.98% (range 0-69%) and 0.18% (range 0-0.87%) in samples from term and preterm groups respectively. All the other seven bacteria together represented only 0.4% of the total reads in the dataset.

Bacteria prevalence (presence/absence) was also assessed (only species with at least 10 total reads were included). A total of 60 taxa had significant differences in prevalence between term and preterm groups; 11 species had greater prevalence in the term cohort and 49 species were more prevalent in the preterm cohort (Table 3.3). *Bifidobacterium infantis*, for example, was two times more prevalent in the term group in comparison with preterm, and *Prevotella timonensis* was 1.58 times more prevalent in the preterm group (Table 3.3). Several *Prevotella* spp. were associated with both term and preterm. *Prevotella amnii* and *P. tannerae* had greater prevalence in the term cohort, whereas *P. timonensis, P. bivia, P. corporis* and *P. bucalis* were more prevalent in the preterm group (Table 3.3).

Mollicutes (*Mycoplasma* and/or *Ureaplasma*) were detected by family-specific conventional PCR in 28/46 (60%) of pregnant women who delivered preterm (Table 3.1). *Ureaplasma* species were detected by genus-specific PCR in samples of 14/46 (30%) women who had PTB, with all women testing positive for *U. parvum* and none for *U. urealyticum*. Women who delivered at term were less likely to be PCR positive for Mollicutes compared to women who had PTB (Table 3.1). No significant differences were observed in *Ureaplasma* prevalence between the two groups (Table 3.1).

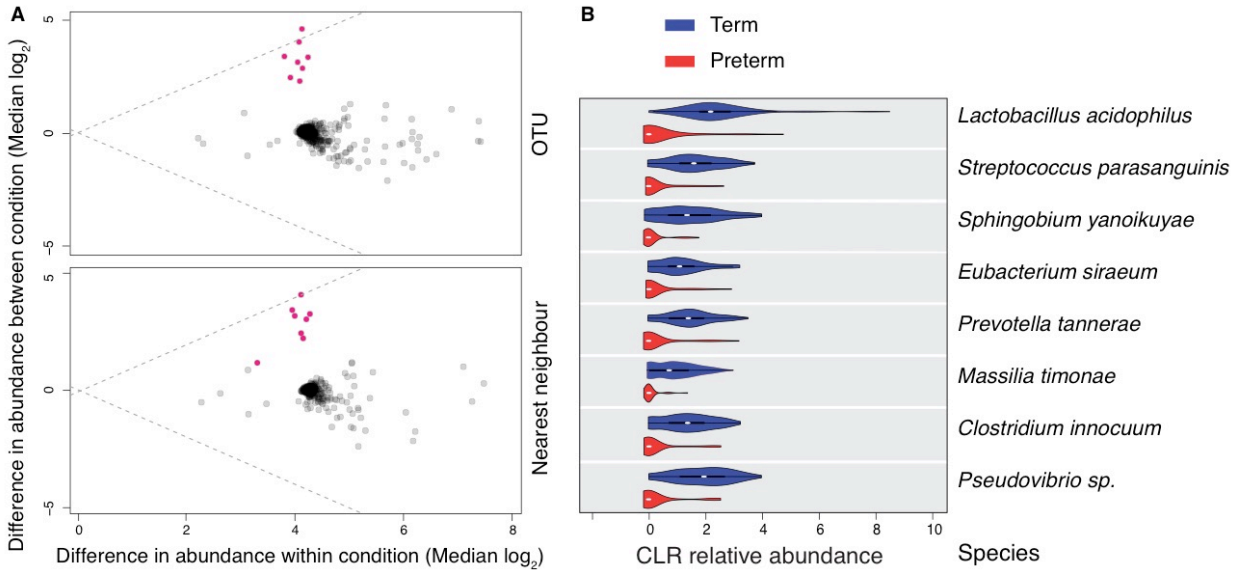

Figure 3.3. Bacteria relative abundance differences between term and preterm groups represented by ALDEx2.

(A) ALDEx2 between- and within-difference values for individual organisms across gestational age category. Organisms (at OTU and nearest neighbour species level) with significant p-values are shown as pink circles (Welch's t statistical test). (B) Violin plots showing the bacteria relative abundance (centre log transformed, CLR) in term and preterm groups. Only the eight bacteria with significant relative abundance differences between term and preterm groups are shown. In the violin plots, the white dot represents the median value, the black bar is the interquartile range, and the vertical width of the plot shows the density of the data along the X-axis

Table 3.3. Bacteria prevalence in the vaginal microbiomes of women who delivered preterm and at term.

Panel A: species with greater prevalence in the term group (ratio term/preterm). Panel B: species with greater prevalence in the preterm group (ratio preterm/term).

Species	Total reads	Preterm (n=46)		Term (n=170)		Prevalence ratio	FDR *
		Reads	Prevalence (%)	Reads	Prevalence (%)		
A						**term/preterm**	
Bifidobacterium infantis	56246	409	19.57	55837	40.00	2.04	0.033
Lactobacillus delbrueckii subsp. *lactis*	48063	9	4.35	48054	28.82	6.63	0.005
Lactobacillus acidophilus	17440	569	89.13	16871	97.65	1.10	0.033
Prevotella amnii	13339	68	4.35	13271	21.76	5.01	0.023
Pseudovibrio sp.	1634	52	15.22	1582	91.76	6.03	0.000
Streptococcus parasanguinis	1199	63	19.57	1136	88.82	4.54	0.000
Sphingobium yanoikuyae	1174	16	13.04	1158	80.00	6.13	0.000
Prevotella tannerae	930	58	15.22	872	88.24	5.80	0.000
Clostridium innocuum	814	42	13.04	772	81.18	6.22	0.000
Eubacterium siraeum	641	43	17.39	598	78.24	4.50	0.000
Massilia timonae	426	6	8.70	420	70.00	8.05	0.000
B						**preterm/term**	
Prevotella timonensis	11450	4211	71.74	7239	45.29	1.58	0.005
Dialister micraerophilus	7381	2850	67.39	4531	44.12	1.53	0.021
Prevotella sp.	2216	626	47.83	1590	14.71	3.25	0.000
Bacteroides coagulans	1283	1212	41.30	71	7.06	5.85	0.000
Corynebacterium accolens	1083	999	52.17	84	13.53	3.86	0.000
Porphyromonas uenonis	1040	795	28.26	245	12.94	2.18	0.038
Actinomyces neuii subsp. *anitratus*	877	738	43.48	139	10.00	4.35	0.000
Peptoniphilus harei	875	811	45.65	64	7.65	5.97	0.000
Lactobacillus fermentum	729	664	17.39	65	5.29	3.29	0.026
Facklamia hominis	711	709	15.22	2	1.18	12.93	0.000
Prevotella bivia	638	228	26.09	410	8.24	3.17	0.005
Prevotella corporis	535	463	23.91	72	5.88	4.07	0.000
Corynebacterium timonense	503	355	34.78	148	10.00	3.48	0.000
Corynebacterium genitalium	372	323	34.78	49	7.06	4.93	0.000
Tepidanaerobacter sp.	329	303	23.91	26	4.12	5.81	0.000
Corynebacterium amycolatum	264	228	32.61	36	4.71	6.93	0.000
Parvimonas micra	193	183	10.87	10	1.18	9.24	0.005
Mobiluncus curtsii subsp. *curtsii*	183	171	21.74	12	2.35	9.24	0.000
Finegoldia magna	155	149	13.04	6	3.53	3.70	0.038
Coprococcus eutactus	154	68	47.83	86	27.06	1.77	0.026
Brevibacterium linens	147	142	10.87	5	0.59	18.48	0.000
Rothia dentocariosa	144	120	30.43	24	5.29	5.75	0.000
Streptococcus thermophilus	132	126	19.57	6	2.35	8.32	0.000
Magnetospirillum magnetotacticum	116	40	43.48	76	24.12	1.80	0.033
Dethiobacter alkaliphilus	101	97	10.87	4	1.76	6.16	0.017
Brevibacterium massiliense	99	86	15.22	13	1.18	12.93	0.000
Eremococcus coleocola	90	82	10.87	8	1.76	6.16	0.017
Anaeromusa acidaminophila	86	85	15.22	1	0.59	25.87	0.000
Arthrobacter globiformis	66	48	10.87	18	1.76	6.16	0.017

Staphylococcus epidermidis	54	47	13.04	7	2.35	5.54	0.009
Corynebacterium simulans	52	41	26.09	11	1.76	14.78	0.000
Anaeroglobus geminatus	52	49	15.22	3	1.18	12.93	0.000
Cellvibrio gilvus	42	26	15.22	16	1.76	8.62	0.000
Prosthecochloris vibrioformis	41	40	6.52	1	0.59	11.09	0.028
Prevotella buccalis	33	33	15.22	0	0.00	-	0.000
Acidaminococcus fermentans	31	29	13.04	2	1.18	11.09	0.000
Brevibacterium casei	25	24	8.70	1	0.59	14.78	0.005
Streptococcus sanguinis	23	18	13.04	5	2.35	5.54	0.009
Atopobium parvulum	23	20	8.70	3	1.18	7.39	0.023
Peptoniphilus duerdenii	22	20	10.87	2	0.59	18.48	0.000
Pelobacter propionicus	21	18	6.52	3	0.59	11.09	0.028
Staphylococcus hominis	21	17	15.22	4	1.76	8.62	0.000
Atopostipes suicloacalis	18	8	10.87	10	1.18	9.24	0.005
Corynebacterium pyruviciproducens	18	13	13.04	5	1.76	7.39	0.005
Corynebacterium jeikeium	14	11	13.04	3	0.59	22.17	0.000
Sporichthya polymorpha	13	11	8.70	2	1.18	7.39	0.023
Rhodococcus jostii	13	7	8.70	6	1.18	7.39	0.023
Nitrospina gracilis	10	5	8.70	5	0.59	14.78	0.005
Megasphaera sp. BV3C16-1	10	9	6.52	1	0.59	11.09	0.028

*FDR (false discovery rate) represents the corrected p-value for multiple comparisons.

Relationships between microbiological and socio-demographic characteristics within the preterm group

The association between CST (I, II, III, IVC, IVD, V) from profiles of women who delivered preterm and several microbiologic-socio-demographic characteristics was investigated. Only two associations were significant: CST and microbiota richness, and CST and microbiota diversity (ANOVA, $p<0.001$). There was no significant association between CST and the remaining following metadata: microbiota richness and diversity (continuous variable), presence of Mollicutes and *Ureaplasma* (yes/no), log 16S rRNA gene copies (continuous variable), maternal age (continuous variable; 18-25, 26-35, 36-45), BMI category (underweight, normal, overweight, obese; <25, ≥25), ethnicity (White, East Asian, South Asian, Black, Hispanic, Other), natural conception (yes/no), parity (0, >1), gestational age (continuous variable), mode of delivery (vaginal, C-section), pre-existing condition (yes/no), folic acid intake before or during pregnancy (yes/no), drinking alcohol (yes/no), neonate in high level care (yes/no), birth weight (continuous variable), and Apgar score at 1 and 5 minutes (1-9).

3.5. Discussion

In this study we determined the composition of the vaginal microbiota of women who had spontaneous preterm birth and compared these profiles to those of women who delivered at term, previously reported by our research group (Freitas et al., 2017). The cohorts were comparable in terms of maternal age, BMI, ethnicity, consumption of tobacco, alcohol and probiotics, which is of interest given that several of these characteristics have been previously associated with preterm delivery. In particular, previously described factors included: low and high maternal ages (Fraser et al., 1995; Jacobsson et al., 2004; Schempf et al., 2007), low BMI (Han et al., 2011), black ethnicity (Kistka et al., 2007), high levels of tobacco, alcohol and illicit drugs consumption (Behrman and Butler, 2007), close temporal proximity to a previous delivery (Conde-Agudelo et al., 2006), and multiple gestation (Stock and Norman, 2010). This cohort is unique in that it did offer the opportunity to have gestational age at delivery as the main characteristic distinguishing these two groups recognizing that the majority of preterm births occurred beyond 32 weeks gestation.

Other known risk factors for sPTB includes maternal medical disorders like hypertension, asthma, diabetes, and thyroid disease (Behrman and Butler, 2007). Although some women in both cohorts reported these conditions, there were not enough participants to stratify the data based on the individual disorder and therefore was not possible to investigate the interaction between those medical conditions and gestation outcome. We were, however, able to confirm previous reports of history of prematurity as a risk factor for preterm birth (Mercer et al., 1999).

Since many organisms isolated from the amniotic cavity of women who experienced preterm birth are also found in the genital tract (Gardella et al., 2004; Krohn et al., 1995; Hillier et al., 1988; Romero et al., 1989), an intrauterine infection ascending from the vagina is one of the currently hypothesized triggers of PTB (Goldenberg et al., 2000). In this study, however, we did not identify a signature microbiota composition (CST) associated with preterm birth. This observation is consistent with the results presented by others (Hyman et al., 2014; Romero et al., 2014b). CST assignments are largely driven by the dominance of a single species, which may mask differences in rare taxa that would differentiate term and preterm groups, and indeed, further analysis revealed that the vaginal microbiota of women who experienced preterm birth was richer and more diverse than those of women who delivered at term. Also, most women in our study were considered late preterm and although we cannot address this question, it is possible that sPTB driven by an ascending infection would be more evident in a high-risk cohort or extreme preterm cases. A recent study of a high-risk pregnant cohort has reported that *L. iners* was strongly associated with short cervix and preterm birth, as *L. crispatus* was associated with term deliveries (Kindinger et al., 2017). Those differences in study outcomes indicate that the pathogenesis of sPTB in low- and high-risk groups might be different.

One controversy that challenges the current hypothesis of preterm caused by an ascending infection is that antibiotic administration to pregnant women with a disturbed vaginal microbiota does not improve outcome in most cases, as demonstrated by study trials (Kekki et al., 2001; Andrews et al., 2006a) and systematic reviews (Morency and Bujold, 2007; Okun et al., 2005; Thinkhamrop et al., 2015). One explanation for the inefficacy of antibiotic treatment in the prevention of preterm birth relies is the high rates of antibiotic resistance among bacterial-vaginosis associated bacteria (Beigi et al., 2004; Budd et al., 2016). In this case, antibiotics not only do not kill the targeted bacteria, but might also reduce the vaginal lactobacilli population leading to an even more disturbed microbiota.

In addition to differences in richness and diversity, differences in the microbiota between the two cohorts regarding bacteria abundance and prevalence were also identified. The ALDEx2 analysis indicated that eight rare taxa were more abundant in the term group, which does not necessarily mean they are associated with a "healthier" state or implicated in preventing sPTB. Since these bacteria are detected at very low abundance within the microbiota profiles, their biological significance in the vaginal microbiome is questionable. Differences in the prevalence of several other taxa between groups were also observed. For example, more women in the term group had *Prevotella amnii* and *P. tannerae* detected in their vaginal samples, whereas *P. timonensis, P. bivia, P. corporis* and *P. buccalis* were more frequently detected in samples from women in the preterm group (Table 3.3). *Prevotella* spp. have been previously associated with bacterial vaginosis and preterm labour (Hill, 1993; Paramel Jayaprakash et al., 2016; Ling et al., 2010), and our results indicate that different *Prevotella* species might have different roles in sPTB. Several of the taxa that were significantly different in their prevalence among women in the two groups also had low sequence read counts (Table 3.3). Further investigation would be required to determine if these rare members of the microbial community play a yet unknown role in sPTB.

It is also important to note that the number of bacterial species with greater prevalence in the preterm (49/60) was higher than in the term (11/60) cohort (Table 3.3), which is consistent with our results of increased microbial richness and diversity in the samples from women who experienced preterm birth. This might indicate that increased richness, rather than the presence of specific taxa, might be associated with sPTB. Those differences might also be an indicator of physiological/biochemical dissimilarities in the vaginal microbiomes of women who deliver at term or preterm. In other words, the physiological state that leads to sPTB might also create an environment that supports a richer/more diverse microbiota.

Our results also confirmed previous reports of an association between *Mycoplasma* and preterm birth (Foxman et al., 2014). Mollicutes were detected significantly more often in women in the preterm group compared to women in the term group, but no differences were observed in *Ureaplasma* prevalence between groups indicating that the difference in Mollicutes prevalence is primarily driven by the presence of *Mycoplasma* spp. Although individual *Mycoplasma* species could not be discerned based on assays used in our study, both *Mycoplasma genitalium* (Hitti et al., 2010; Edwards et al., 2006; Lis et al., 2015) and *Mycoplasma hominis* (Donders et al., 2009;

Kwak et al., 2014; Hillier et al., 1995; Lamont et al., 1987) have been previously associated with negative reproductive outcomes including PTB.

Collectively, our overall findings were similar to other two studies, which provided us the opportunity to compare different study designs that addressed the same research question. Hyman and colleagues (Hyman et al., 2014) described the vaginal microbiota of 83 pregnant women (term n=66, preterm n=17) based on Sanger sequencing of cloned 16S rRNA. Samples were collected at each trimester and preterm was defined as delivery before 37 weeks of gestation. There was no correlation between preterm and absence/low abundance of *Lactobacillus* in the microbiota; in other words, preterm outcome could not be predicted based on CST. Similar to our results, they found an association between increased microbiota diversity and preterm delivery among women of white ethnicity (n=40) (data from women of other ethnicities was not included in the analysis because of small sample sizes).

Romero and colleagues (Romero et al., 2014b) also investigated the vaginal microbiota of pregnant women who experienced preterm delivery, defined as delivery before 34 weeks of gestation (term n=72, preterm n=18). The profiles were created by 16S rRNA amplicon sequencing, and samples were collected every 4 weeks until 24 weeks of gestation and then every 2 weeks. They found no differences in the frequency of different CST between women who had term and preterm deliveries. Likewise, no differences in bacterial relative abundance were observed between the two cohorts, although only bacteria that were present in at least 25% of samples were included in the analysis. These results are consistent with our findings of bacterial abundance based on the ALDEx analysis since we only found significant differences in relative abundance for eight rare bacteria. Unlike Hyman et al. (Hyman et al., 2014) and our results, Romero et al. (Romero et al., 2014b) did not find differences in microbiota diversity between women who delivered preterm and at term. One possible explanation for this contradictory result might be related to differences in participant ethnicity among these studies. While most women in our study and the Hyman et al. study described themselves as white, the majority of participants in the Romero et al. study described themselves as African American. It has been reported that the composition of the vaginal microbiota is strongly associated with a woman's ethnicity (Ravel et al., 2011; Zhou et al., 2007). Other studies have also demonstrated that black ethnicity is associated with an increased microbiota diversity in comparison with white ethnicity (Fettweis et al., 2014),

which could have masked differences in bacterial diversity between term and preterm cohorts in the Romero study.

Contrary to our overall findings, DiGiulio and colleagues (DiGiulio et al., 2015) found a strong association between the non-*Lactobacillus*-dominated CST IV and preterm birth in a case-control study based on the 16S rRNA amplicon sequencing. Pregnant women (preterm n=34, term n=15), mostly of white ethnicity, were sampled weekly throughout gestation. Interestingly, the authors pointed out that if samples had been collected less frequently, short-term "excursions" to CST IV would have been missed and probably the association between CST IV and preterm birth would have been less obvious. The detection of a temporary microbiota disturbance represented by a change from a *Lactobacillus*-dominated CST to CST IV may have been missed in our study since samples were not collected longitudinally.

Most samples in the preterm group were dominated by *Lactobacillus*, yet, they collectively had higher richness and diversity compared to samples from the term group. The increased microbiota richness/diversity might indicate a transient state between *Lactobacillus*-dominated CST and non-*Lactobacillus*-dominated, i.e., CST IV (A, C or D). In other words, the increased richness and diversity we observed might be a remnant characteristic of the previous disturbed microbiota. In summary, although we did not "detect" a specific microbial community structure that is associated with preterm birth, the increased microbiota richness/diversity might be an indication that the bacterial community overwhelmed placental and amniotic defences and permitted ascending bacterial contamination of the chorioamnion, which could still lead to preterm birth. In addition, the association with differences in types of *Prevotella* species and *Mycoplasma* presence may point to signature species associated with preterm birth.

3.6. Conclusions

Taken together, our results suggest that the differences in the microbiota of women who had preterm deliveries, such as increased microbiota richness and diversity and greater prevalence of Mollicutes and other bacteria, may have a role in sPTB. Other differences between cohorts might have been masked by the presence of highly dominant bacteria like *Lactobacillus*. At the overall level, we did not identify a specific vaginal microbial community structure at 11-16 weeks gestation age that predicts sPTB. Also, differences in relative abundance of bacterial species

between term and preterm groups were only significant for a few low abundance species. Although a causal relationship remains to be determined, our results confirm previous reports of an association between Mollicutes and preterm birth, and further suggest that a diverse bacterial community may contribute to the microbiome's role in sPTB. Alternatively, the more rich and diverse microbiotas of the preterm group may reflect physiological differences between the groups that affect selection of bacteria. This study provides valuable evidence of subtle alterations in the microbiome associated with preterm birth that requires further study utilizing sequencing methodology. In addition, future study should include evaluation of the microbial metabolite production and host response might further elucidate factors leading to sPTB and identify women at risk early in pregnancy.

Acknowledgements

The authors are grateful to the women who participated in the study and acknowledge the contribution and support of Ontario Birth Study Team members and the participants. Financial support was provided by a joint Canadian Institutes of Health Research (CIHR) Emerging Team Grant and a Genome British Columbia (GBC) grant awarded to DMM, AB and JEH (grant reference #108030) as well as a CIHR grant MOP-82799 to AB. ACF was supported by a University of Saskatchewan graduate scholarship. Funding for the Ontario Birth Study was provided by Mount Sinai Hospital Foundation, Lunenfeld-Tanenbaum Research Institute and the Department of Obstetrics and Gynecology at Mount Sinai Hospital.

Table 3.4S. Supplementary material (.xls). Summary of OTU analyzed in this study.

OTU ID, percentage of identity, length, cpnDB name, species, and abundance in each library are shown.

CHAPTER 4. Quantification, isolation and characterization of *Bifidobacterium* from the vaginal microbiomes of reproductive aged women

Citation

Freitas AC & Hill JE (2017). Quantification, isolation and characterization of *Bifidobacterium* from the vaginal microbiomes of reproductive aged women. *Anaerobe*, 47, 145–156. doi.org/10.1016/j.anaerobe.2017.05.012

Authors contributions

Conceived and designed the experiments: ACF and JEH. Performed the experiments and analyzed the data: ACF. Wrote the paper: ACF and JEH.

Chapter transition

Several studies have investigated the causes of microbiome dysbiosis with limited success, regardless of pregnancy status. Researchers have also tried to understand the association between the microbiota and preterm birth, but still there is no consensus regarding which bacteria within the microbiome, if any, specifically lead to preterm birth. One difficulty in studying such complex biological system is the limited and poor definition of "healthy" and "unhealthy" microbiome. The current definition of a healthy microbiome is to be dominated by lactobacilli, which is not observed in many clinically healthy women. To better understand the relationship among the microbiota, host and reproductive outcomes, an improved definition of healthy microbiome is needed. In this chapter I described several vaginal microbial profiles in "healthy" women dominated by bifidobacteria. Bifidobacteria are a common member of the gut microbiome and have been associated with several health promoting effects in the gastrointestinal tract. The first goal was to confirm that bifidobacteria are the dominant bacteria in the microbiome of some women. My second objective was to characterize vaginal bifidobacteria isolates with a focus on phenotypic characteristics common to lactobacilli, and associated with a protective microbiome. Improving the understanding the role of these organisms in the vaginal microbial community contributes to an improved definition of "healthy", which focuses on the phenotypic properties of bacteria and not only in the presence/absence of specific taxa.

4.1. Abstract

The vaginal microbiome plays an important role in women's reproductive health. Imbalances in this microbiota, such as the poorly defined condition of bacterial vaginosis, are associated with increased susceptibility to sexually transmitted infections and negative reproductive outcomes. Currently, a "healthy" vaginal microbiota in reproductive aged women is understood to be dominated by *Lactobacillus*, although "atypical" microbiomes, such as *Bifidobacterium*-dominated profiles, have been described. Despite these observations, vaginal bifidobacteria remain relatively poorly characterized and questions remain regarding their actual abundance in the microbiome. In this study, we used quantitative PCR to confirm the relative abundance of *Bifidobacterium* in the vaginal microbiomes of healthy reproductive aged women, previously determined by deep sequencing. We also isolated and phenotypically characterized vaginal bifidobacteria (n=40) in the context of features thought to promote reproductive health. Most isolates were identified as *B. breve* or *B. longum* based on cpn60 barcode sequencing. Fermentation patterns of vaginal bifidobacteria did not differ substantially from corresponding type strains of gut or oral origin. Lactic acid was produced by all vaginal isolates, with *B. longum* strains producing the highest levels, but only 32% of isolates produced hydrogen peroxide. Most vaginal bifidobacteria were also able to tolerate high levels of lactic acid (100 mM) and low pH (4.5 or 3.9), conditions typical of vaginal fluid of healthy women. Most isolates were resistant to metronidazole but susceptible to clindamycin, the two most common antibiotics used to treat vaginal dysbiosis. These findings demonstrate that *Bifidobacterium* is the dominant member of some vaginal microbiomes and suggest that bifidobacteria have the potential to be as protective as lactobacilli according to the current understanding of a healthy vaginal microbiome.

Keywords: *Bifidobacterium*, vaginal microbiome, fermentation pattern, hydrogen peroxide, lactic acid, antibiotic.

4.2. Introduction

Bifidobacteria were first described by Tissier in 1899, who isolated a bacterium from breast-fed infant feces and named it *Bacillus bifidus* (Poupard et al., 1973). In 1924, Orla-Jensen proposed the genus *Bifidobacterium* as a separate taxon for these organisms (Orla-Jensen, 1924), which currently includes more than 30 species (Biavati and Mattarelli, 2012). Bifidobacteria are

Gram-positive, anaerobic, non-motile, non-spore forming rod-shaped bacteria, with varied branching. They belong to the *Bifidobacteriaceae* family and have high genomic G+C content (55-67 mol%) (Biavati and Mattarelli, 2012). Bifidobacteria are known to colonize the human vagina, oral cavity and, more abundantly, the gastrointestinal tract (GIT) (Biavati and Mattarelli, 2006).

Several studies have shown their influence on human physiology and nutrition (Arboleya et al., 2011; Fukuda et al., 2011; Martinez et al., 2013; Sugahara et al., 2015; Meyer and Stasse-Wolthuis, 2009). In newborns, bifidobacteria play an important role as one of the primary colonizers of the GIT, representing 60 to 91% of fecal bacteria in breast-fed infants (Harmsen et al., 2000; Turroni et al., 2012). This proportion decreases with age and it may represent less than 10% of the adult fecal microbiota (Yatsunenko et al., 2012; Turroni et al., 2008). Bifidobacteria provide protection from pathogens in the GIT through the production of bacteriocins (Martinez et al., 2013), inhibition of pathogen adhesion (Arboleya et al., 2011) and modulation of the immune system (Hart et al., 2004; Fanning et al., 2012). Due to these health-promoting effects, bifidobacteria have been extensively studied as probiotics (Sugahara et al., 2015; Zinedine and Faid, 2007; Chenoll et al., 2011; Heuvelin et al., 2010).

Early microbial colonization is an essential process in the maturation of the immune system (Rautava et al., 2012). This initial colonization may be affected by many factors, such as the mode of delivery (vaginal or caesarean section), feeding type (breast-fed or formula-fed), exposure to antibiotics and hygiene (Fanaro et al., 2003). However, the relative contributions of maternal microbiota (gut, breast milk, vaginal) and environmental sources to the bifidobacteria population of the neonatal gut remain unresolved.

While *Bifidobacterium* spp. present in the gut are well described, vaginal bifidobacteria remain relatively poorly characterized, and it is not known if vaginal adaptation has resulted in distinct phenotypic features that distinguish them from gut populations. Although a healthy vaginal microbiota is defined as *Lactobacillus*-dominated, several studies have identified vaginal *Bifidobacterium*-dominated profiles in 5-10% of healthy, reproductive aged women (Chaban et al., 2014; Hyman et al., 2005; Shipitsyna et al., 2013; Sundquist et al., 2007). Furthermore, some vaginal bifidobacteria are able to produce lactic acid and hydrogen peroxide; attributes of vaginal lactobacilli credited with maintaining homeostasis in the vaginal microbiome (Schellenberg et al., 2012).

Culture-independent techniques are useful tools in microbiome characterization, but methods based on amplification and sequencing of 16S rRNA genes, have been reported to underrepresent *Bifidobacterium* in microbial communities (Hill et al., 2010). The abundance of *Bifidobacterium* in the vaginal microbiota may also be underestimated due to the similarity of their 16S rRNA sequences to those of *Gardnerella vaginalis*. *G. vaginalis* is also a member of the *Bifidobacteriaceae* family and is a commonly detected microorganism associated with bacterial vaginosis (BV) (Burton et al., 2003b). The use of the cpn60 "universal target" (UT) region as a barcode for microbiome profiling results in better resolution of closely related species, including those within *Bifidobacteriaceae* (Links et al., 2012), and cpn60 based human fecal microbiome profiles have been shown to more accurately represent *Bifidobacterium* content than a 16S rRNA based approach (Hill et al., 2010). Previous studies of the vaginal microbiome (Chaban et al., 2014) or synthetic mixtures of vaginal organisms (Schellenberg et al., 2017) have demonstrated a strong correlation between cpn60 sequence read abundance and organism abundance determined by quantitative PCR. However, regardless of the target used, relative abundance of specific organisms within complex communities may not be represented accurately by methods that rely on polymerase chain reaction (PCR) amplification and its inherent biases.

Considering the lack of information about *Bifidobacterium* spp. of vaginal origin, their importance as a potential source for the neonatal gut microbiome, and their potential health-promoting effects in the vagina, a better understanding of the properties of vaginal bifidobacteria is needed. In this study, our main objectives were: 1) to apply species-specific quantitative PCR to confirm the relative abundance of *Bifidobacterium* in the vaginal microbiomes of reproductive aged women previously determined based on high-throughput DNA sequencing methods, and 2) to characterize vaginal *Bifidobacterium* isolates based on carbohydrate fermentation patterns, hydrogen peroxide production, lactic acid production, resistance to low pH and lactic acid, and susceptibility to antibiotics.

4.3. Methods

Samples and microbiome profiles

Vaginal microbiome profiles from 492 healthy women were previously published by our research group (Albert et al., 2015; Freitas et al., 2017). Profiles were created by PCR amplification

and deep sequencing of the cpn60 UT region. Total bacterial load in each sample was also estimated as part of these studies using a SYBR Green assay based on the amplification of the V3 region of the 16S rRNA gene. The remaining vaginal swabs and DNA extracts from these studies, archived at -80 ˚C, were available for use in the current study.

Bifidobacterium quantitative PCR assays

Sequences with similarity to *Bifidobacterium breve*, *Bifidobacterium dentium* and *Alloscardovia omnicolens* (*Bifidobacteriaceae* family) that were detected at high frequency in the previously published studies (Albert et al., 2015; Freitas et al., 2017) were selected as targets for quantitative PCR. Signature regions within the cpn60 UT unique to each target were determined using Signature Oligo software (LifeIntel Inc., Port Moody, BC, Canada) and primers were designed using Primer-blast software (Ye et al., 2012) and Primer3 (Rozen and Skaletsky, 2000) (Table 4.1).

Table 4.1. Primers used in the bifidobacteria qPCR assays.

Target	Primer name	Sequence (5'->3')	Product size (bp)	Annealing Temperature (°C)
B. breve	JH0472 (F)	AACCGTGCTCGCCCAGTC	150	65
	JH0473 (R)	TCCTTGGTCTCAACGTCCTT		
B. dentium	JH0474 (F)	GTGCTCGAAGACCCGTACAT	163	65
	JH0475 (R)	GGATGGTGTTCAGGATCAGG		
A. omnicolens	JH0470 (F)	GCACGAAGGCTTGAAGAACG	197	65
	JH0471 (R)	CCAAAGCCTCAGCAATACGC		

To create plasmids for use in standard curves, target sequences were amplified from vaginal swab DNA extracts. The resulting PCR products were purified and ligated into cloning vector pGEM-T-Easy (Promega, Madison, WI) and used to transform competent E. coli DH5α. Insertion of the intended target sequence was confirmed by DNA sequencing. Optimal annealing temperature for each assay was determined using an annealing temperature gradient, and specificity of each primer set was confirmed by using plasmids containing cpn60 UT sequences from closely related species as template.

All qPCR reactions were performed in duplicate and each batch of reactions included a no template control and a standard curve consisting of serial dilutions of plasmids containing targets. Each reaction consisted of 2 µL of template DNA, 1× iQ SYBR Green Supermix (BioRad, Mississauga, ON, Canada) and 400 nM each primer, in a final volume of 25 µL. A MyiQ thermocycler (BioRad) was used for all reactions with the following protocol: 95 °C for 3 m, followed by 40 cycles of 95 °C for 15 s, 65 °C for 15 s, 72 °C for 15 s. A dissociation curve was subsequently performed for 81 cycles at 0.5 °C increments from 55 °C to 95 °C to confirm the purity of PCR products.

Calculation of proportional abundance of *Bifidobacterium* in vaginal samples

Vaginal microbiome profiles (n=492) were ranked according to the proportional abundance of *Bifidobacterium* (all *Bifidobacterium* species and *Alloscardovia* combined), based on the previously determined cpn60 sequence read counts (Albert et al., 2015; Freitas et al., 2017). *B. breve*, *B. dentium* and *A. omnicolens* DNA was quantified in vaginal swab DNA extracts from selected samples using the SYBR Green qPCR assays described in the previous section. Previously determined total 16S rRNA copy number per sample (Albert et al., 2015; Freitas et al., 2017) was used an as estimate of total bacterial population. The ratio between $\log_{10}$ copy number of *Bifidobacterium* and 16S rRNA $\log_{10}$ copy number was used as an estimate of the proportional abundance of each target species in the selected vaginal microbiome samples.

Proportional abundance determined from deep sequencing of cpn60 UT amplicons (percent of sequence reads) and by quantitative PCR for each of the three targets evaluated (*B. breve*, *B. dentium* and *A. omnicolens*) were compared using Spearman rank correlation in IBM SPSS (Statistical Package for the Social Sciences, version 21).

Isolation of vaginal *Bifidobacterium*

A complete list of isolates used in the study and their sources is provided in Table 4.2. Bifidobacteria were isolated from vaginal swabs, which were collected in previous studies and had been stored at -80 °C. For one group of samples (healthy pregnant and non-pregnant Canadian women from Vancouver, BC and Toronto, ON), 12 *Bifidobacterium*-dominated samples were selected based on the previous microbial profiling data (Albert et al., 2015; Freitas et al., 2017). Eluted material from these 12 vaginal swabs was plated on Columbia agar containing 5% sheep blood (CSB, BD Canada, Mississauga, ON). For the second group of samples (Adolescent women, Winnipeg, MB) (Schellenberg et al., 2012), material from 27 vaginal swabs was plated on a Bifidus selective medium agar (BSM agar, Sigma-Aldrich, Oakville, ON); no information about the microbiome was available. Frozen swabs were thawed and sample was eluted from each swab by vortexing in phosphate-buffered saline (PBS). Dilutions were prepared from the eluted sample and spread onto CSB agar or BSM agar followed by incubation using the GasPak EZ anaerobic system (BD Canada, Mississauga, ON) at 37 °C for 72 h. After isolating pure colonies, a freezing buffer (4% (v/v) skim milk, 1% (w/v) glucose, 20% (v/v) glycerol) was added to the isolates for long-term storage at -80 °C. DNA preparations from isolates were made using Chelex 100 (Bio-Rad Laboratories, Inc., Mississauga, ON).

Bifidobacteria (n=16) previously isolated in a study of commercial sex workers in Nairobi, Kenya (Schellenberg et al., 2012), were also included in the study. *Bifidobacterium* spp. and *Lactobacillus crispatus* type strains (*B. breve* ATCC 15700, *B. longum* subsp. *infantis* ATCC 15697, *B. dentium* ATCC 27534 and *L. crispatus* ATCC 33820) were acquired from the American Type Culture Collection (Manassas, VA). *Lactobacillus crispatus* vaginal isolates 67-1, B12-1 and N4D05 were from our lab culture collection.

Table 4.2. Complete list of vaginal isolates used in this study.

cpnDB nearest type strain match	Isolate ID	ID (%)	Origin
Bifidobacterium breve	(I)30-1	99.5	Canada
	(I)30-2	99.5	Canada
	(IV)30-1A	99.5	Canada
	(I)91-1	98.9	Canada
	(I)91-2	99.1	Canada
	(I)91-3	99.1	Canada
	(I)322-1	99.5	Canada
	(I)322-2	99.5	Canada
	(IV)322	99.5	Canada
	12-4	99.8	Canada
	(W)20-13	99.3	Canada
	(W)20-15	99.3	Canada
	(W)20-17	99.3	Canada
	(W)56	99.3	Canada
	N6D12	99.1	Kenya
Bifidobacterium longum subsp. *longum*	(I)239-2	99.3	Canada
	(I)239-4	99.3	Canada
	(IV)239	99.1	Canada
	(W)35-1	100.0	Canada
	N2E12	100.0	Kenya
	N2F05	100.0	Kenya
	N2G10	100.0	Kenya
	N3A01	100.0	Kenya
	N3E01-2	100.0	Kenya
	N5E04	100.0	Kenya
	N6D05	100.0	Kenya
Bifidobacterium dentium	(VI)131-1	99.5	Canada
	(W)90-1	99.1	Canada
	N3E11	99.5	Kenya
	N5E10	99.1	Kenya
Bifidobacterium bifidum	(W)15	100.0	Canada
	(W)27-1	100.0	Canada
	(W)86-2	100.0	Canada
Bifidobacterium adolescentis	N1D05	95.1	Kenya
	N5F04	94.4	Kenya
Bifidobacterium kashiwanohense	N4G05	97.8	Kenya
	N5G01	97.6	Kenya
Bifidobacterium catenulatum	N3F01	99.1	Kenya
Bifidobacterium pseudocatenulatum	N4E05	99.3	Kenya
Alloscardovia omnicolens	(IV)256-6	99.5	Canada

cpn60 PCR, sequencing and phylogenetic analysis

The cpn60 UT region was used for species identification of isolates. The target region was amplified from genomic DNA using modified versions of the cpn60 "universal" primers, optimized for high G+C templates: H1594 (5'-CGC CAG GGT TTT CCC AGT CAC GAC GAC GTC GCC GGT GAC GGC ACC ACC AC-3') and H1595 (5'-AGC GGA TAA CAA TTT CAC ACA GGA CGA CGG TCG CCG AAG CCC GGG GCC TT-3'). The M13 (-40)F and M13 (-48)R sequencing primer landing sites are underlined.

PCR was carried out with 2 µL template DNA in a reaction mixture containing 1× PCR buffer (0.2 M Tris-HCl at pH 8.4, 0.5 M KCl), 2.5 mM $MgCl_2$, 200 µM dNTP mixture, 400 nM each primer, 2 U AccuStart Taq DNA polymerase and water to a final reaction volume of 50 µL. Cycling conditions were as follows: 94 °C for 3 m, 40 cycles of 94 °C for 30 s, 57 °C for 60 s, 72 °C for 60 s, and a final extension at 72 °C for 10 m. Amplification was confirmed by resolving the PCR products in 1% agarose gel (expected cpn60 UT amplicon is ~652 bp). Purified PCR products were sequenced with primers M13(-40)F or M13(-48)R. Following assembly of the forward and reverse sequences and removal of amplification primer sequences, the resulting 552 bp sequences were compared to the cpnDB reference database (www.cpndb.ca) (Hill et al., 2004) for identification.

Phylogenetic analysis of isolates was done using PHYLIP (Phylogeny Inference Package) version 3.5 (Felsenstein, 1993). The cpn60 sequences of the isolates and reference strains from the cpnDB reference database were aligned with ClustalW (Thompson et al., 1994). Alignments were bootstrapped with seqboot; distances were calculated with the maximum likelihood option of dnadist. Dendrograms were constructed from distance data by neighbor-joining with neighbor. Consensus trees were calculated with consense, and branch lengths were superimposed on consensus trees using dnaml.

Carbohydrate fermentation patterns

The ability of vaginal *Bifidobacterium* isolates to metabolize 49 different carbohydrates was tested using the API 50 CH kit and API 50 CHL broth (bioMérieux, France) according to manufacturer's instructions. Assay strips were incubated anaerobically at 37 °C for 8 days. Type

strains *B. breve* ATCC 15700, *B. longum* subsp. *infantis* ATCC 15697 and *B. dentium* ATCC 27534 were included in the assay for comparison with vaginal isolates. Previously published descriptions of carbon source utilization by other type strains were also used for comparison (*B. longum* subsp. *longum* (Reuter, 1971), *B. bifidum* (Scardovi et al., 1971), *B. adolescentis* (Reuter, 1971), *B. kashiwanohense* (Morita et al., 2011), *B. catenulatum* (Scardovi and Crociani, 1974), *B. pseudocatenulatum* (Scardovi et al., 1979) and *A. omnicolens* (Huys et al., 2007)).

Results for each isolate were recorded as positive or negative for acid production from each carbon source based on colour reaction. Acid production (colour change) after 2 days of incubation was considered a delayed reaction. A Jaccard index was used to measure pairwise distances between isolates based on their fermentation patterns by using the vegan package (function 'vegdist') in R (Oksanen et al., 2015). This distance matrix was used for hierarchical clustering using the 'hclust' function in R, with UPGMA (Unweighted Pair Group Method with Arithmetic Mean).

Hydrogen peroxide assay

A modified Brucella agar media (mB) containing 0.25% (w/v) tetramethylbenzidine (TMB, Sigma-Aldrich, Oakville, ON) and 0.01% (w/v) horseradish peroxidase (HRP, Sigma-Aldrich, Oakville, ON) was used for chromogenic detection of hydrogen peroxide (H_2O_2) production (Rabe and Hillier, 2003). Isolates grown on CSB agar were streaked onto mB agar plates. Alternatively, 1 mL of sterile PBS was used to wash colonies off the CSB plate and the resulting suspension was adjusted to 1 McFarland and spotted on to mB plates using a sterile replicator. mB plates were incubated anaerobically at 37 °C for 72 h, removed from the incubator and held at room temperature in air for 30 minutes for colour development. Hydrogen peroxide production was reported as positive or negative based on the colour of the colonies (no colour/pale beige = negative, green/blue = positive).

Lactic acid assay

Isolates were grown overnight in Modified Reinforced Clostridial broth (for *Bifidobacterium* and *Alloscardovia*) or deMan, Rogosa and Sharpe broth (MRS) (for

Lactobacillus). All cultures were adjusted to an initial optical density (OD_{600}) ≈ 0.1. After incubation, broth cultures were pelleted and supernatants were heat-treated at 80 °C for 15 minutes to stop enzymatic reactions. Lactic acid quantification was conducted using the D- and L-lactic acid Enzymatic BioAnalysis/Food Analysis UV method kit (R-Biopharm, Darmstadt, Germany). Results were reported as concentration of lactic acid per OD_{600}. Each experiment was repeated at least three times. Concentration of lactic acid produced by different species was compared by Kruskal-Wallis test using IBM SPSS (Statistical Package for the Social Sciences, version 21).

Tolerance of bifidobacteria to low pH and lactic acid

Experiments to test tolerance of bifidobacteria to lactic acid and low pH were based on O'Hanlon, Moench and Cone (2011) (O'Hanlon et al., 2011) with minor modifications. A subset of *Bifidobacterium* spp. (n=18) and *Gardnerella vaginalis* (n=3) isolates were grown overnight in Modified Reinforced Clostridial broth (pH 6.8) or NYC III broth (pH 7.3), respectively. An aliquot (50 µL) of the overnight inoculum was added to each control or experimental medium (final volume 5 mL) and then incubated anaerobically at 37 °C for 2 hours. Control media was prepared at pH 6.8. Experimental media was prepared at pH 4.5 with the following concentrations of lactic acid: 0 mM, 1 mM, 10 mM, 100 mM and 1000 mM, and at pH 3.9 with (100 mM) and without lactic acid (0 mM). After 2 hours exposure, each sample was serially diluted with the appropriate medium containing 200 mM HEPES and track-plated (Jett et al., 1997). The pH of each experimental or control medium was re-measured after the 2 hours incubation to confirm it had remained within 0.2 pH units of the starting pH. Agar plates were incubated anaerobically at 37 °C for 48 h and colony forming units (cfu/mL) were counted. Each experiment was repeated at least three times.

The percent bacterial survival was calculated based on the ratio of colony counts (log_{10} cfu/mL) of bacteria that had been incubated in experimental media (treatment) compared to colony counts (log_{10} cfu/mL) of bacteria that had been incubated in control media (control), according to the following formula:

$$Percent\ survival = \frac{log_{10}\,{}^{cfu}/_{mL}\,treatment}{log_{10}\,{}^{cfu}/_{mL}\,control} \times 100$$

Susceptibility to antibiotics

Susceptibility of bifidobacteria to clindamycin and metronidazole was evaluated using Etest strips (bioMérieux, France). Overnight cultures grown in Modified Reinforced Clostridial broth were adjusted to 1 McFarland. The suspension was spread evenly on Columbia sheep blood agar plates and Etest strips were placed on the agar surface. Plates were incubated anaerobically at 37 °C for 48 h. The minimum inhibitory concentration (MIC) for each antibiotic was read as the lowest antibiotic concentration at which growth was inhibited.

4.4. Results

Confirmation of *Bifidobacterium*-dominated vaginal microbial profiles

We identified 21/492 (4.2%) of the previously published vaginal microbiome profiles that were dominated (>50% of sequence reads) by *Bifidobacterium*-like sequences, of which eight were dominated by *B. breve*, five by *B. longum*, three by *B. dentium* and five by *A. omnicolens*. An additional 6% (29/492) of microbiome profiles had intermediate (1-50%) levels of *Bifidobacterium*-like sequences and 59.5% (293/492) of profiles had low (<1%) levels. *Bifidobacterium*-like sequences were undetected in 30.3% (149/492) of samples.

All samples (n=492) were ranked based on the total *Bifidobacterium*-like relative abundance, previously determined by cpn60 amplicon sequencing. A total of forty-two samples with high (>50%, n=11), medium (1-50%, n=11), low levels (<1%, n=10) and undetected *Bifidobacterium*-like sequences (n=10) (i.e., samples with ≤ 2 reads) were randomly selected from the 492 samples for qPCR analysis. This sample set (n=42) was assayed for each of the three targets (*B. breve*, *B. dentium* and *A. omnicolens*) by species-specific qPCR. The copy number for each target determined by qPCR was expressed as a proportion of the total bacterial load, estimated by 16S rRNA gene copy number (Albert et al., 2015; Freitas et al., 2017). Sequence read numbers (% abundance), qPCR results for *B. breve*, *B. dentium* and *A. omnicolens*, and previously determined total 16S rRNA copy numbers are provided in Supplementary Table 4.4S, and results are summarized in Figure 4.1.

For samples in the high category, the three targets analyzed (*B. breve*, *B. dentium* and *A. omnicolens*) comprised approximately 100% of the estimated total bacterial load (Figure 4.1, left

panel). In two of the *B. breve* dominated samples where *A. omnicolens* was also detected in the microbiome profiles, both targets were detected by qPCR. *B. breve* and *B. dentium* were also detected by qPCR in two of the three high category samples where the microbiome profiles were dominated by other *Bifidobacterium* species not targeted by any of the qPCR assays. All five samples with medium levels (1-50% of sequence reads) of *B. breve* based on cpn60 amplicon sequencing, had detectable *B. breve* sequences by qPCR (Figure 4.1, middle panel). Estimates of proportional abundance of *B. breve* in these samples based on qPCR (77-113%), however, were much higher than the proportional abundance of corresponding sequence reads in the cpn60 sequence-based microbiome profiles (all <10%). Two of the three samples with medium levels of *B. dentium* were qPCR positive for this target, but again the qPCR based estimate of proportional abundance (82-102%) was much higher than that predicted by sequence read numbers (both <25%). Samples with medium levels of *A. omnicolens* were negative by species-specific qPCR, but *A. omnicolens* was detected by qPCR in other samples in the medium category that were dominated by *B. breve*, *B. dentium* or other bifidobacteria. Samples in the low and undetected categories were negative for all three qPCR assays (Figure 4.1, right panel).

When all samples (n=42) were considered, regardless of abundance category, there was a positive correlation between cpn60 amplicon sequencing data (% of sequence reads) and qPCR values ($\log_{10}$ copies per swab) for all three targets (*B. breve* ρ=0.671, $p<0.0001$; *B. dentium* ρ=0.502, $p=0.001$; *A. omnicolens* ρ=0.784, $p<0.0001$). The correlations remained significant when samples with zero values were removed from analysis.

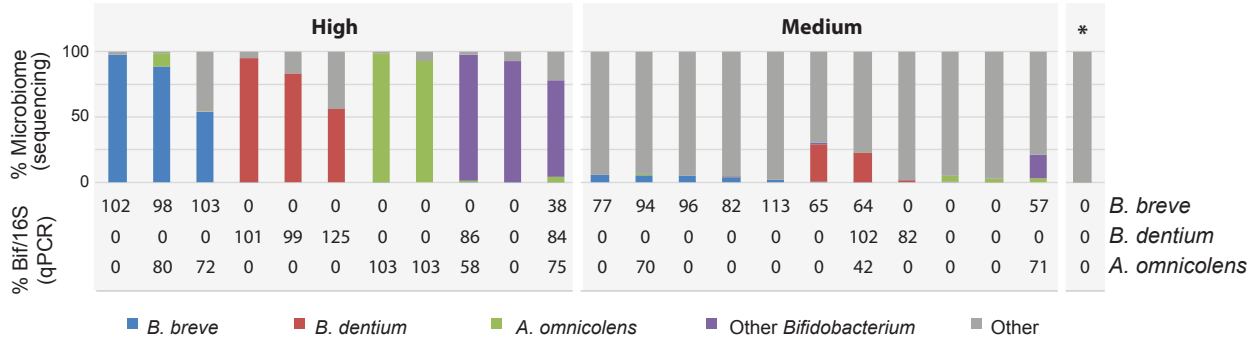

Figure 4.1. Detection of *B. breve*, *B. dentium* and *A. omnicolens* by sequencing and qPCR, in 42 vaginal microbiomes.

Results are shown for samples in the high (n=11, left panel), medium (n=11, middle panel) and low/undetected (n=20, right panel with *) bifidobacteria abundance categories as determined by *cpn*60 amplicon sequencing. All samples in the "low" (n=10) and "undetected" (n=10) categories had identical results so only one example is shown. The bar charts illustrate the proportion of sequences assigned to each species detected, indicated by colour according to the legend. The lower table shows the percentage of *Bifidobacterium* (*B. breve*, *B. dentium* or *A. omnicolens* $\log_{10}$ copies) out of the total bacterial load (estimated by 16S rRNA gene $\log_{10}$ copies) (%Bif/16S (qPCR)).

Vaginal *Bifidobacterium* isolates

A total of 40 isolates from the vaginal swabs of 26 women from Canada and Kenya were included in the study (Table 4.2). Based on the cpn60 gene PCR amplification, sequences of the 40 isolates were compared to cpnDB and the nearest type strain sequence was identified as *B. breve* n=15 (98.9-99.8% sequence identity over 552 bp), *B. longum* n=11 (99.1-100%), *B. dentium* n=4 (99.1-99.5%), *B. bifidum* n=3 (100%), *B. adolescentis* n=2 (94.4-95.1%), *B. kashiwanohense* n=2 (97.6-97.8%), *B. catenulatum* n=1 (99.1%), *B. pseudocatenulatum* n=1 (99.3%) and *Alloscardovia omnicolens* n=1 (99.5%). A phylogenetic tree based on the cpn60 UT sequences of the 40 vaginal isolates and 12 reference sequences from type strains of human origin is shown in Figure 4.2. *B. adolescentis* and *B. kashiwanohense* isolates clustered separately from their respective type strains with good bootstrap support.

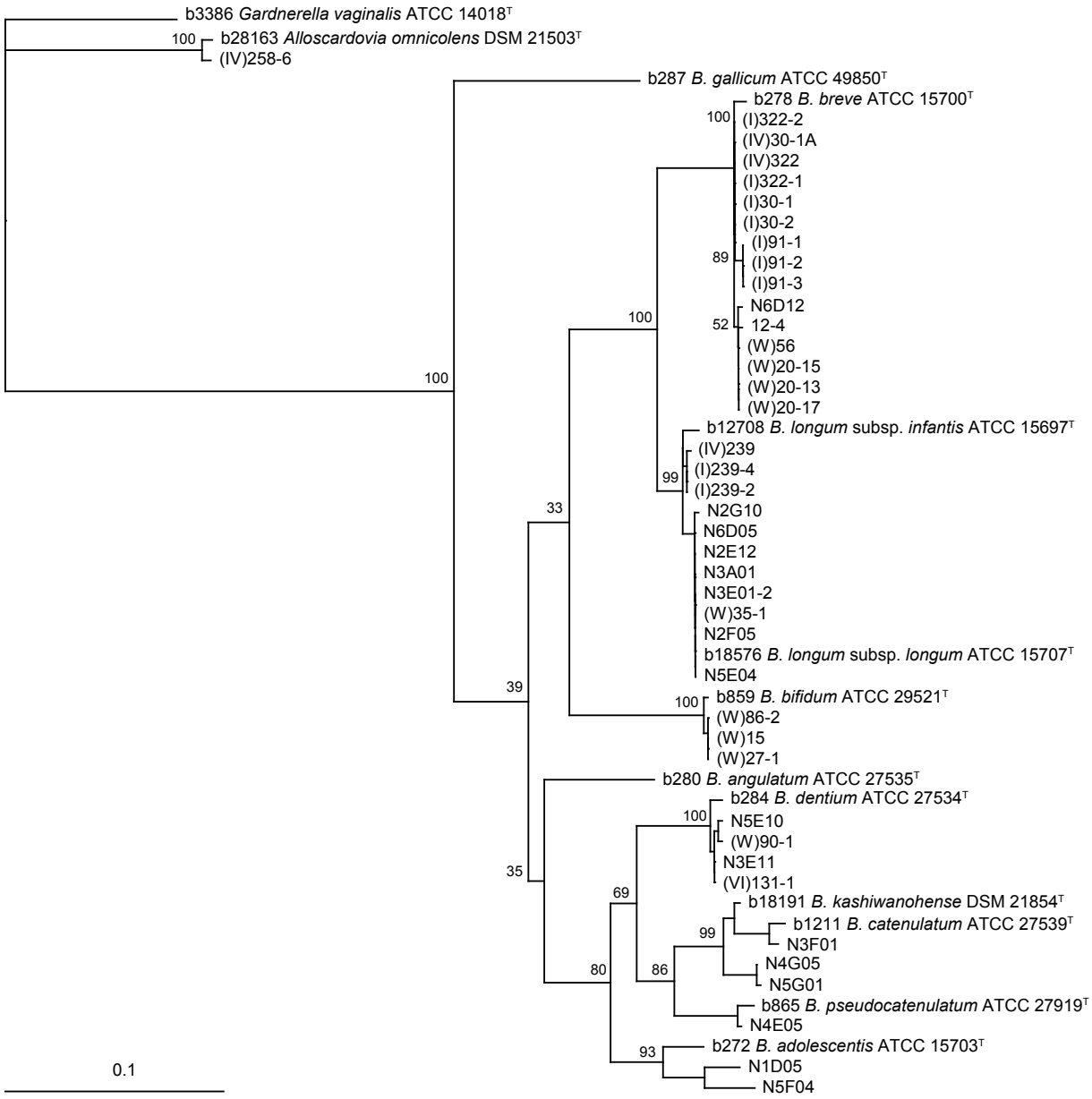

Figure 4.2. Phylogenetic tree based on cpn60 UT sequences of vaginal bifidobacteria and reference strains[T].

The tree was rooted with *Gardnerella vaginalis* ATCC 14018T and constructed using the Dnaml method with bootstrap values calculated from 100 trees. The number at each node represents the percentage bootstrap support.

Carbohydrate fermentation patterns

Vaginal bifidobacteria isolates were tested for their ability to metabolize 49 different carbon sources. The relatedness of the fermentation patterns of vaginal bifidobacteria (n=39) and three available type strains (n=3) was visualized by UPGMA clustering (Figure 4.3). As expected, most strains clustered based on species identity, except for the two *B. kashiwanohense*-like isolates (N4G05 and N5G01) and three *B. longum* isolates ((W)35-1, (I)239-2 and (IV)239). (I)239-2 and (IV)239 were the only *B. longum* isolates that metabolized glycogen, gentiobiose and L-fucose, and (W)35-1 was the only *B. longum* strain not able to metabolize D-ribose. Also, most *B. longum* isolates (7/11) did not cluster with the type strain (*B. longum* subsp. *infantis* ATCC 15697) based on their fermentation patterns. The use of trehalose by *B. breve* differed between the type strain and most vaginal isolates (type strain was negative and 13/15 vaginal isolates were positive). Vaginal *B. dentium* isolates mainly differed from the type strain due to their ability to use D-cellobiose (4/4). Overall, for *B. breve* isolates (n=15), there was a complete agreement between vaginal isolates and the type strain ATCC 15700 (gut origin) for utilization of 71% (35/49) of carbon sources tested. For *B. longum* (n=10), vaginal isolates and type strain ATCC 15697 (gut origin) had a complete agreement for 51% (25/49) of carbon sources. For *B. dentium* (n=4), a complete agreement of 84% (41/49) was observed between vaginal isolates and type strain ATCC 27534 (oral cavity origin). Overall, there were no remarkable differences in the fermentation patterns between vaginal bifidobacteria and their respective type strains of gut or oral cavity origin.

A second analysis was performed considering the percentage of vaginal isolates that were positive for each carbon source in comparison to the literature description of the type strains (Supplementary Table 4.5S). This analysis enabled the comparison of isolates for which the type strain was not available for inclusion in our fermentation assays. For *B. adolescentis*, vaginal isolates (2/2) did not utilize D-sorbitol and D-melezitose, unlike the type strain (Scardovi et al., 1971). *B. kashiwanohense* isolates (2/2) differed from their type strain due to the ability to metabolize D-trehalose and D-melezitose (Morita et al., 2011). Regarding *B. bifidum*, the main difference was the fermentation of D-raffinose by all vaginal isolates (3/3), which differs from the type strain and other gut *B. bifidum* strains (Biavati and Mattarelli, 2012). Also, vaginal *B. bifidum* did not ferment D-cellobiose, diverging from its type strain (Biavati and Mattarelli, 2012). Despite these few differences, overall utilization patterns of vaginal isolates did not differ substantially from literature descriptions of the type strains.

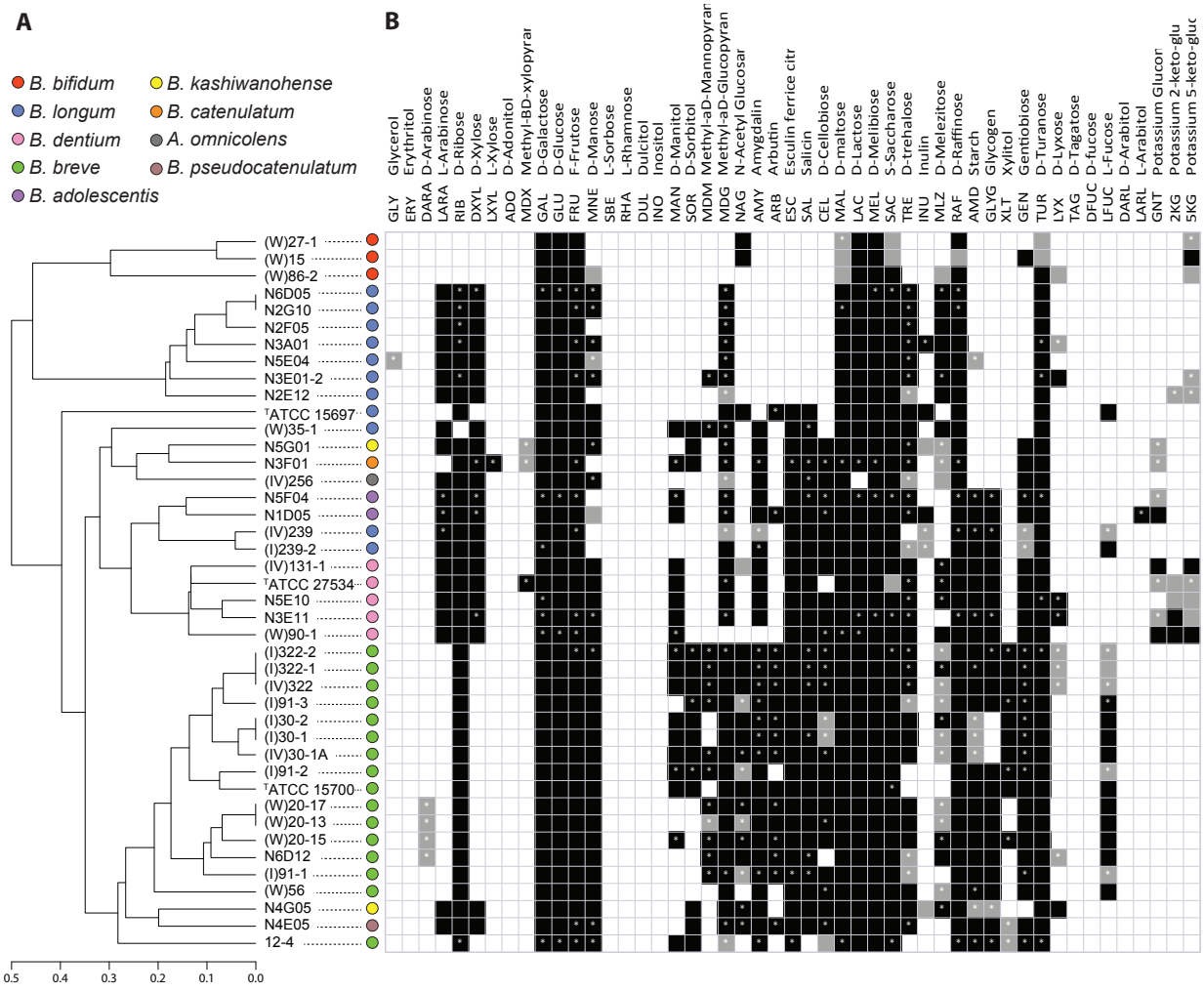

Figure 4.3. Carbohydrate fermentation pattern.

(A) UPGMA dendrogram derived from Jaccard's similarity coefficients calculated among isolates based on their fermentation patterns (isolates n=39, type strainsT n=3). (B) Heatmap representing the fermentation of 49 carbon sources. Black = positive reaction; Grey = weak reaction; White = no growth/fermentation; * = Delayed reaction.

Production of hydrogen peroxide

Hydrogen peroxide production was detected in 32.5% (13/40) of the vaginal bifidobacteria. All *B. dentium* (n=4) strains were positive and all *B. kashiwanohense* (n=2), *B. catenulatum* (n=1), *B. pseudocatenulatum* (n=1) and *A. omnicolens* (n=1) were negative. Production of hydrogen peroxide by *B. breve* (3/15), *B. longum* (4/11), *B. bifidum* (1/3) and *B. adolescentis* (1/2) was variable (Figure 4.4).

Production of lactic acid

Lactic acid production was measured for all vaginal bifidobacteria isolates using a commercial assay. Four vaginal isolates of *L. crispatus*, a species associated with high lactic acid production in the vagina, were included for comparison. All vaginal bifidobacteria produced lactic acid (n=40) (Figure 4.5). While bifidobacteria produced L-lactic acid only, *L. crispatus* produced both D and L lactic acid isomers. As expected, there was significant correlation between the absolute values of total lactic acid concentration (mM) and supernatant pH ($p < 0.01$, Pearson = 1), and culture OD_{600} ($p < 0.0001$, Pearson = 1) (data not shown). A Kruskal-Wallis test was conducted to evaluate differences in the production of lactic acid among species (only species with two or more isolates were included in the statistical analysis). The overall test was significant ($\chi 2$ (6, n=30) = 24.7, $p \leq 0.0001$), hence pairwise comparisons among the groups were conducted with Mann-Whitney test. The highest concentration of lactic acid among the bifidobacteria was produced by *B. longum* isolates, which did not differ from the *L. crispatus* strains (Mann-Whitney, p=0.240) (Figure 4.5).

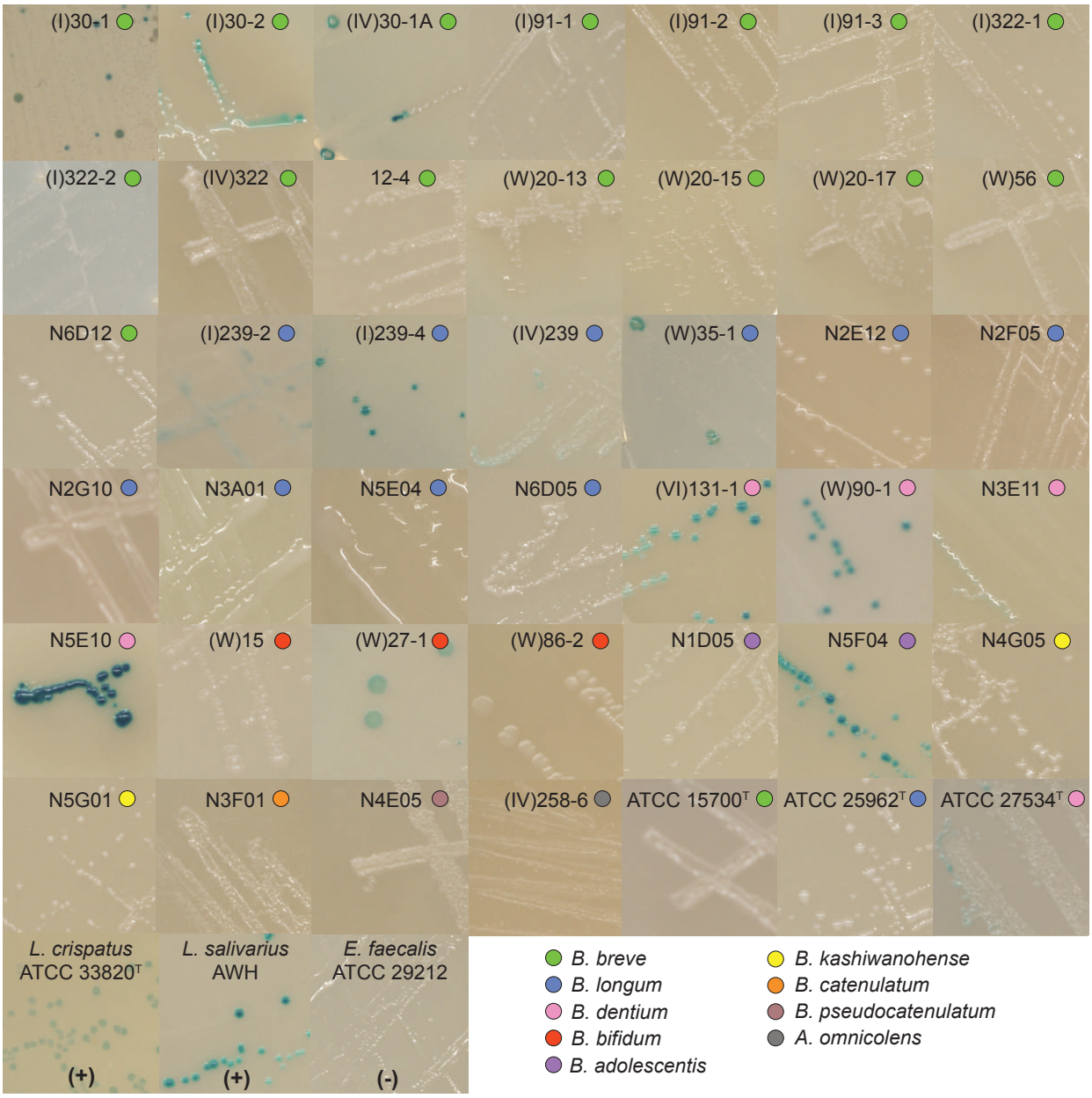

Figure 4.4. Hydrogen peroxide production indicated by blue colour on TMB medium.

Species designation of each isolate is indicated by coloured dot according to the legend. Positive controls (+). Negative control (-).

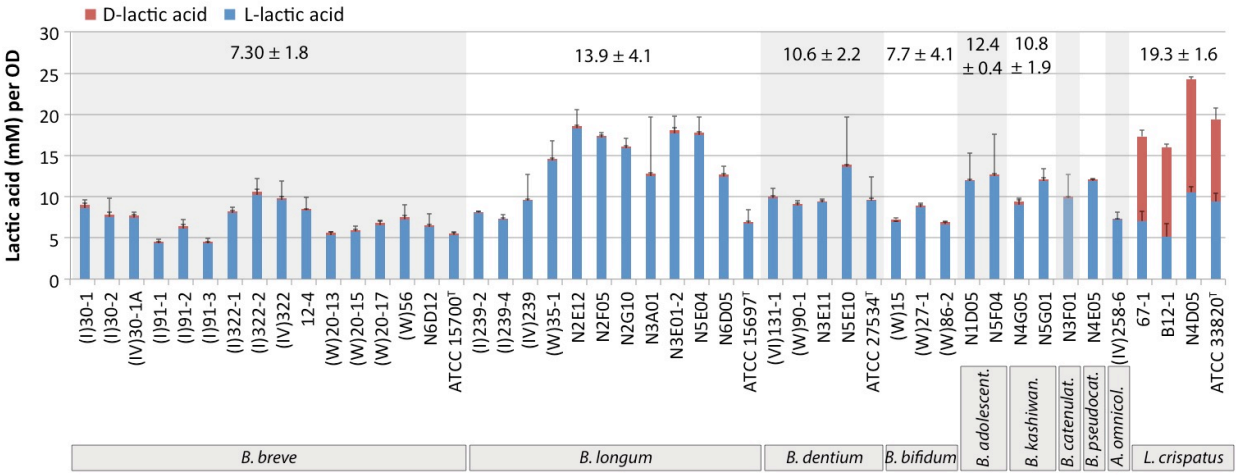

Figure 4.5. Total lactic acid concentration in culture supernatant of vaginal isolates and type strains.

Grey boxes below the chart indicate species identification for each isolate. Numbers at the top are the average lactic acid concentrations (mM/OD_{600} ± standard deviation) produced by each species group (excluding type strains^{T}).

Tolerance to low pH and lactic acid

A subset of vaginal bifidobacteria isolates (15/40) and three type strains were selected for testing tolerance of low pH and high lactic acid concentrations and the percent survival values for each isolate under different conditions are shown in Figure 4.6. All bifidobacteria tested were able to survive at pH 4.5 with up to 100 mM lactic acid; conditions typical of vaginal fluid of healthy women with a *Lactobacillus* dominated microbiome and that have been shown to be lethal to many vaginosis associated bacteria (O'Hanlon et al., 2011). Survival of a few isolates was affected by 1000 mM lactic acid, which is ten times more concentrated than physiologic levels. Surprisingly, pH 4.5 and 100 mM lactic acid did not affect *G. vaginalis*. The experiment was repeated using media at pH 3.9 (pKa of lactic acid), which resulted in drastic reduction in *G. vaginalis* survival, regardless of the inclusion of lactic acid. Most bifidobacteria were tolerant to this more acidic condition, with or without lactic acid (Figure 4.6). *A. omnicolens* and *B. kashiwanohense* (n=2) were the isolates most affected by lactic acid, with survival declining considerably when incubated with 100 mM lactic acid (pH 3.9). Two bifidobacteria type strains, *B. breve* ATCC 15700 and *B. longum* subsp. *infantis* ATCC 15697, were also affected by this condition, with only 45% survival compared to the culture incubated at neutral pH.

Susceptibility to antibiotics

Selected representative isolates were tested for susceptibility to clindamycin and metronidazole, the two antibiotics most widely used to treat vaginal dysbiosis (Table 4.3). Only one isolate (*B. breve* (I)30-1) presented a MIC higher than 8 μg/mL for clindamycin, which is the breakpoint of resistance for anaerobic bacteria according to The National Committee for Clinical Laboratory Standards (CLSI, 2014). For metronidazole, 15/22 isolates had a MIC higher than ≥ 32 μg/mL, the breakpoint for metronidazole resistance (CLSI, 2014). All fifteen of these isolates had a MIC ≥ 256 μg/mL, the maximum antibiotic concentration in the Etest strip.

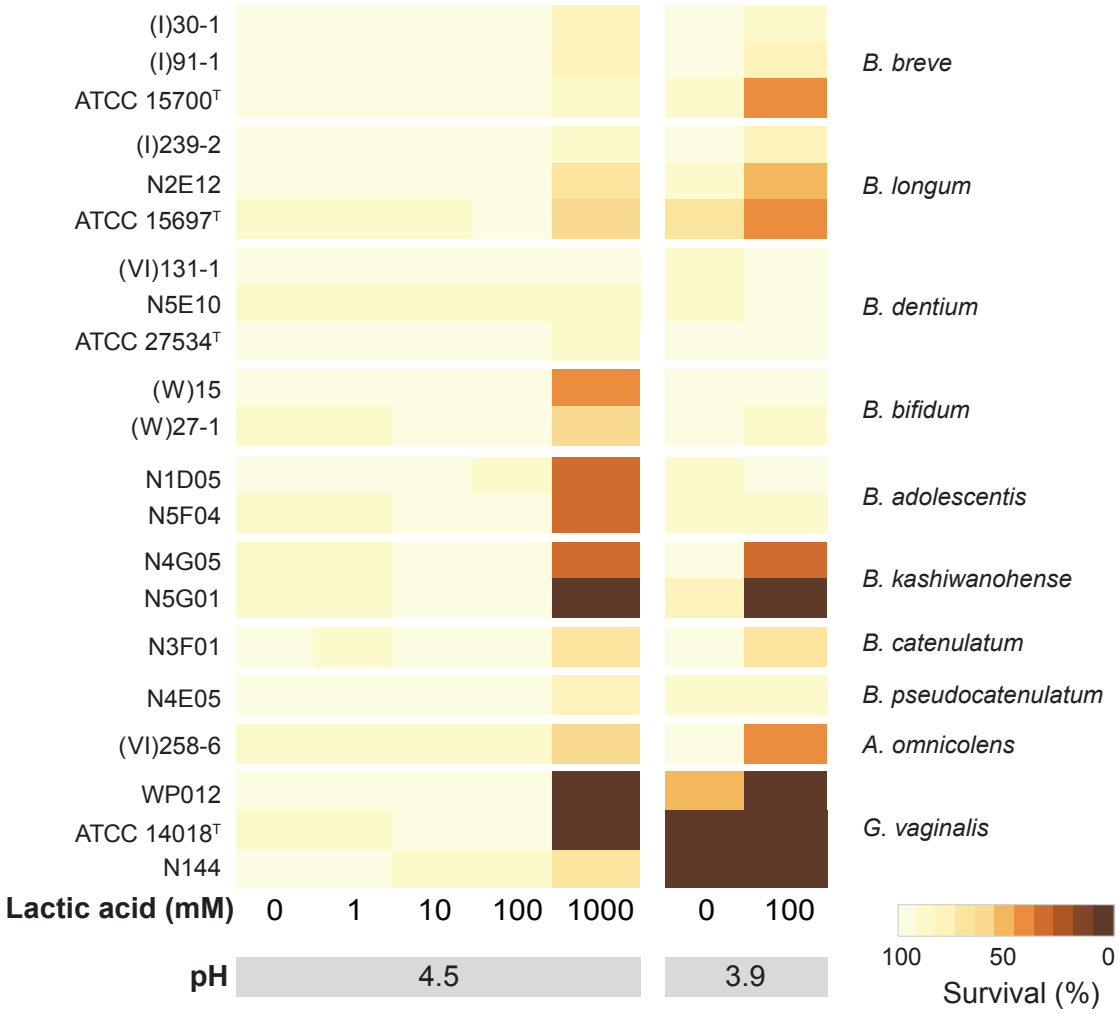

Figure 4.6. Survival (% log cfu/mL) of vaginal bifidobacteria after incubation at low pH and high lactic acid concentrations in comparison with bacteria cultured at pH 6.8.

Results shown are the average of at least three replicate experiments.

Table 4.3. Susceptibility of bacteria to metronidazole and clindamycin.

		Minimum inhibitory concentration (MIC µg/mL)	
Species	**isolate**	**Metronidazole**	**Clindamycin**
B. breve	(I)30-1	8.0	32.0
	(I)91-1	>256	≤0.016
	(I)322-1	>256	0.035
	(W)56	>256	0.032
B. longum	(I)239-2	>256	≤0.016
	N2E12	>256	0.056
	(W)35-1	2.0	0.047
	N2F05	>256	0.125
B. dentium	(VI)131-1	4.0	≤0.016
	N3E11	>256	0.250
	N5E10	4.0	≤0.016
	(W)90-1	>256	≤0.016
B. bifidum	(W)15	2.0	0.032
	(W)27-1	2.0	0.032
	(W)86-2	4.0	0.056
B. adolescentis	N1D05	>256	0.158
	N5F04	>256	0.028
B. kashiwanohense	N4G05	>256	0.047
	N5G01	>256	0.028
B. catenulatum	N3F01	>256	0.040
B. pseudocatenulatum	N4E05	>256	0.040
A. omnicolens	(VI)258-6	>256	0.035

4.5. Discussion

As a result of microbiome characterization by culture-independent, DNA sequence based methods there is a growing appreciation of "atypical" vaginal microbiomes in healthy women, such as the *Bifidobacterium*-dominated profiles. However, it is known that DNA extraction methods and PCR amplification biases affect the sequencing outcome, resulting in a view of the abundances of species within the community that is inevitably distorted to some degree. Thus, for a careful investigation of the microbial composition, further evaluation using alternative techniques is recommended. Bifidobacteria present a particularly interesting and important subject for this type of investigation due to their recognized importance in the neonatal gut, their probiotic potential, and their frequent underrepresentation in PCR based microbiome profiles. Furthermore, the current diagnostic definition of a healthy vaginal microbiome is limited to a *Lactobacillus*-dominated community, and the available gold standard diagnostic method (Gram stain and Nugent scoring (Nugent et al., 1991)) is interpreted based on this restrictive definition.

In this study, we used species-specific qPCR to quantify bifidobacteria DNA in vaginal samples showing a wide range of proportional abundances of bifidobacteria based on cpn60 amplicon sequencing. Our results showed a strong positive correlation between sequence read abundance and qPCR values for the three species tested (Supplementary Table 4.4S). However, when we expressed bifidobacteria abundance as a proportion of the estimated total bacterial population, the values calculated for the medium abundance category (1-50% bifidobacteria sequences in the cpn60 microbiome profile), estimates were much higher than expected. One possible explanation for this is that the 16S rRNA based estimates were biased (underestimated) due to the composition of the samples, resulting in a corresponding overestimation of proportional abundance of bifidobacteria DNA. The biases of "universal" 16S rRNA gene primers are well recognized (Walker et al., 2015; Brooks et al., 2015), making estimates of total bacterial population size challenging using this approach. Better approaches include flow cytometry (Schellenberg et al., 2008), which counts bacterial cells rather than relying on detection of sequences. However, fresh samples are preferred for this approach, which were not available for our study. While it is also possible that the cpn60 amplicon sequencing method underestimated the abundance of bifidobacteria due to preferential amplification of other sequences present in the samples, this seems unlikely given the abundance of *Bifidobacterium* sequence reads in other libraries within the same studies (i.e. the high category) and previous demonstrations of the efficient amplification

of *Bifidobacterium* cpn60 sequences from complex samples (Hill et al., 2010). A few samples with medium levels of bifidobacteria sequence reads were qPCR negative for targets expected to be present. The medium category included samples containing a wide range (1-50%) of bifidobacteria abundance, and the lack of qPCR detection in samples with bifidobacteria levels near the 1% cut off may be due to the detection limit of the assays. Taken together, our results confirm that cpn60 amplicon sequencing reflects relative abundance of bifidobacteria and that *Bifidobacterium* spp. can be the dominant component of the vaginal microbiome. Future studies using more accurate methods for total bacterial population quantification will be needed to better describe this relationship in the context of overall community size.

To learn more about the characteristics of vaginal bifidobacteria detected in sequence-based microbiome studies, we collected 40 isolates from vaginal swabs of clinically healthy women, with most isolates being identified as *B. breve* or *B. longum* (Figure 4.2). A few differences were observed in carbon source utilization patterns between vaginal isolates and type strains, however, most of those differences have also been described among other bifidobacteria of gut or oral cavity origin. For example, unlike the type strain, vaginal *B. bifidum* isolates did not ferment D-cellobiose, but this characteristic is also absent in some gut *B. bifidum* (Biavati and Mattarelli, 2012). Similarly, our vaginal *B. dentium* isolates and other oral cavity isolates utilize D-cellobiose, which is not done by the type strain (Biavati and Mattarelli, 2012). Thus, this lack of agreement in carbon source utilization between vaginal isolates and type strains does not necessarily indicate that there is a distinctive fermentation profile of vaginal bifidobacteria. Our results support the idea that different body sites, such as gut and vagina, host similar strains of bifidobacteria. However, a more detailed examination of the utilization of nutrients that differentiate these environments and the genome contents of vaginal and gut isolates of *Bifidobacterium* spp. may yet reveal signs of niche specialization that are not apparent in examinations of housekeeping functions.

Hydrogen peroxide and lactic acid production have been associated for decades with the protective role of vaginal *Lactobacillus*, but more recent findings have called into question the significance of the physiological concentrations of H_2O_2 produced by *Lactobacillus*. Vaginal H_2O_2-producing lactobacilli are claimed to inhibit growth of opportunist bacteria, which share the same niche with lactobacilli and may not have protective mechanisms like production of catalase or peroxidase (Strus et al., 2006). However, it is unclear whether *Lactobacillus* is able to produce

H_2O_2 under the hypoxic condition of the vagina (Hill et al., 2005a), since oxygen is required for its production. Additionally, cervicovaginal fluid and semen contain proteins, glycoproteins, polysaccharides, and lipids that react with and inactivate H_2O_2 (O'Hanlon et al., 2010). O'Hanlon, Moench, & Cone (2011) also reported that physiological concentrations of H_2O_2 had no microbicidal activity, while a supraphysiologic concentration of H_2O_2 that was sufficient to inactivate bacterial vaginosis associated bacteria, also inactivated vaginal lactobacilli. We decided to test our bifidobacteria isolates for H_2O_2 production considering the wealth of previous studies and descriptions of H_2O_2 production for *Lactobacillus*, the presumed indicator of a healthy vaginal microbiome. Our results indicate that in vitro H_2O_2 production is not widespread in vaginal bifidobacteria (Figure 4.4).

Lactic acid is a key element in promoting a healthy vaginal environment by preventing the overgrowth of bacterial vaginosis-associated microorganisms (O'Hanlon et al., 2011). One of the protective mechanisms attributed to lactic acid is lowering the vaginal pH (Boskey et al., 2001, 1999). It has also been demonstrated that lactic acid is able to disrupt the outer membrane of Gram negative bacteria (Alakomi et al., 2000). Bacteria are known as the primary source of lactic acid in the vagina, with some species being able to produce both lactic acid enantiomers (D and L), as opposed to human cells that only make the L form (Boskey et al., 2001). Here, we confirmed that vaginal *Bifidobacterium* produced only L-lactic acid, as previously described for other bifidobacteria (Kandler, 1983). While *L. crispatus* and *L. gasseri* are both able to produce a mixture of DL-lactic acid, other *Lactobacillus* spp. have been shown to produce only L- (*L. iners*) or D-lactic acid (*L. jensenii*) (Witkin et al., 2013). At physiological concentrations (~56 to 111 mM), both forms of lactic acid are effective in decreasing HIV infectivity in vitro; at lower concentrations, L-lactic acid has greater virucidal activity than D-lactic acid (Aldunate et al., 2013). L-lactic acid has also been shown to have greater antibacterial effect against *E. coli* (O157 and non-O157) than D-lactic acid (Leitich et al., 2003). Although the biological significance of producing different amounts of D and L forms remains unclear, the results of these studies suggest that the microbicidal effect of lactic acid involves more than acidity alone. Since bifidobacteria in general are known lactic acid producing bacteria, our observation of its production is far from surprising. However, what is noteworthy is that the levels of lactic acid produced by vaginal bifidobacteria, especially *B. longum*, are comparable to those of *L. crispatus*, the organism most often associated with a healthy vaginal microbiome (Petrova et al., 2015).

Considering that lactic acid is a hallmark of a healthy vagina, knowing the levels of acid produced by different bifidobacteria is of interest when evaluating whether a vaginal microbiome dominated by *Bifidobacterium* should be considered as protective as a microbiome dominated by *Lactobacillus*. Moreover, assessing the ability of bifidobacteria to persist in high levels of lactic acid and low pH, conditions typical of the vaginal fluid of healthy women, was an important step in improving knowledge of bifidobacteria ecology. Other studies have shown that protonated lactic acid rather than lactate anion is the microbicidal form of lactic acid (O'Hanlon et al., 2011; Aldunate et al., 2013), which emphasizes the role of pH in mediating lactic acid activity. We initially conducted our experiments at pH 4.5, recreating conditions used in previous investigations of microbicidal activity of lactic acid against vaginosis-associated bacteria (O'Hanlon et al., 2011). However, these conditions did not kill any of the three *G. vaginalis* isolates included in our study. We repeated the experiments at pH 3.9, a value of which is now understood to be within the range of expected pH for the cervicovaginal fluid of clinically healthy women (~3.5 - 4.5) (O'Hanlon et al., 2013) and observed differential survival of *G. vaginalis* and the bifidobacteria. Bifidobacteria resistance to low pH and high concentrations of lactic acid may be species specific, as isolates from the same species behaved similarly in most cases (Figure 4.6). *B. kashiwanohense*-like isolates, for example, were clearly less tolerant to these acidic conditions than *B. breve*, *B. dentium*, *B. bifidum* or *B. adolescentis* isolates. Further studies of larger numbers of isolates will be required to test this observation.

The normal microbiota plays an important role in the recovery of the microbiome after antibiotic treatment. While most vaginal bifidobacteria were susceptible to clindamycin, we observed high rates of metronidazole resistance. Metronidazole and clindamycin are the two antibiotics recommended by CDC to treat bacterial vaginosis, and can be administrated either orally or intravaginally (CDC - Centers for Disease Control and Prevention, 2015b). Although metronidazole is the first-line treatment against bacterial vaginosis (Hay, 2009), many non–spore-forming, Gram-positive anaerobic rods are resistant to it, including *Propionibacterium, Atopobium, Mobiluncus*, *Bifidobacterium* and *Lactobacillus* (Hall and Copsey, 2015). *Gardnerella vaginalis*, a hallmark microorganism in BV, and *Prevotella* have also been demonstrated to be resistant to metronidazole (Alauzet et al., 2010; Aroutcheva et al., 2001). Since *G. vaginalis*, *Atopobium, Mobiluncus* and *Prevotella* are considered vaginosis associated bacteria, this might explain the high rates of BV recurrence after metronidazole treatment (Hay, 2009). Although

antimicrobial resistance is a concern in BV treatment, bifidobacteria resistance might be a beneficial factor by facilitating the microbiota recovery after antibiotic administration.

In this study we confirmed that a subset of healthy, reproductive aged women have vaginal microbiomes dominated by *Bifidobacterium* spp. We also demonstrated that vaginal bifidobacteria have the potential to be as protective as lactobacilli according to the current understanding of a "healthy" vaginal microbiome. These results have significant implications for women's health diagnostics since current protocols based on Gram staining and Nugent score would likely result in a diagnosis of "intermediate" or "consistent with BV" if a vaginal smear was dominated by *Bifidobacterium* rather than *Lactobacillus*. We expect our findings will help to guide clinicians and researchers to better assess a healthy vaginal microbiome, and avoid unnecessary interventions.

Acknowledgments

The authors would like to thank John Schellenberg for providing some of the bifidobacteria isolates and Champika Fernando for insightful discussions during the execution of this study. ACF was supported by a University of Saskatchewan graduate scholarship. The authors are grateful to the Vaginal Microbiome Group Initiative (VOGUE) team members Deborah Money, Alan Bocking, Sean Hemmingsen, Janet Hill, Gregor Reid, Tim Dumonceaux, Gregory Gloor, Matthew Links, Kieran O'Doherty, Patrick Tang, Julianne Van Schalkwyk and Mark Yudin for access to samples, and helpful discussions.

Supplementary material

Table 4.4S. Supplementary material (.xls). Sequence read numbers and qPCR results.

Sequence read numbers (% relative abundance), qPCR results for *B. breve*, *B. dentium* and *A. omnicolens*, and previously determined total 16S rRNA copy numbers of 42 samples used in the screening analysis.

Table 4.5S. Supplementary material (.xls). Carbohydrate fermentation patterns of vaginal bifidobacteria isolates and of their type strain.

Numbers represent the percentage of vaginal isolates that had positive reaction for each carbon source.

Available at:

http://www.sciencedirect.com/science/article/pii/S1075996417301129?via%3Dihub#appsec1

CHAPTER 5. Bifidobacteria isolated from vaginal and gut microbiomes are indistinguishable by comparative genomics

Citation

Freitas AC & Hill JE. Bifidobacteria isolated from vaginal and gut microbiomes are indistinguishable by comparative genomics. (Submitted, November 1st, 2017).

Authors contributions

Conceived and designed the experiments: ACF and JEH. Performed the experiments and analyzed the data: ACF. Wrote the paper: ACF and JEH.

Chapter transition

Bifidobacteria are found in the human gut (adult and infant), vaginal and milk microbiomes, and although they have been the subject of numerous studies due to their gut health-promoting effects, the mechanisms of mother-infant transmission remain in discussion. Considering their potential importance in the vaginal and infant-gut microbiome, I have investigated several aspects of vaginal bifidobacteria in the previous study chapter. First, I confirmed that bifidobacteria can be the dominant bacteria in the vaginal microbiome of a subpopulation of women. Second, I described bifidobacteria characteristics that are consistent with the phenotypic properties of lactobacilli, the "healthy" gold standard bacteria. Although the results indicated that bifidobacteria may play a similar role as lactobacilli, it is unknown if the vaginal bifidobacteria are an adapted subpopulation or if they are ubiquitous bacteria that have health promoting effects in both the gut and the vagina. To elucidate some aspects of these possible adaptations and indirectly infer aspect of mother-infant bifidobacteria transmission, I assessed the genetic variability of *B. breve* and *B. longum* isolated from two different microbiomes: vaginal and gut. These two bifidobacteria species are the most common bifidobacteria species found in the gut (adult and infant) and genital tract. The results of this study demonstrated that gut and vaginal bifidobacteria have very similar genome repertoires indicating they represent the same bacterial population.

5.1. Abstract

Bifidobacteria colonize the human gastrointestinal tract, vagina, oral cavity and breast milk. They influence human physiology and nutrition through health-promoting effects, play an important role as primary colonizers of the newborn gut, and contribute to vaginal microbiome homeostasis by producing lactic acid. Nevertheless, the mechanisms by which bifidobacteria are transmitted from mother to infant remains in discussion. Moreover, studies have suggested that *Bifidobacterium* spp. have specializations for gut colonization, but comparisons of strains of the same bifidobacteria species from different body sites are lacking. Here, our objective was to compare the genomes of *Bifidobacterium breve* (n = 16) and *Bifidobacterium longum* (n = 26) to assess whether gut and vaginal isolates of either species were distinguishable based on genome content. Comparison of the general genome features showed that vaginal and gut isolates did not differ in size, GC content, number of genes and CRISPR, either for *B. breve* or *B. longum*. Average nucleotide identity and whole genome phylogeny analysis revealed that vaginal and gut isolates did not cluster separately. A few differences in the accessory genes between vaginal and gut strains were observed, but were not sufficient to distinguish isolates based on their origin. Vaginal and gut isolates also had a similar COG (Cluster of Orthologous Group) category distribution, and no evidence of differential adaptation to either body site was identified. The results of this study support the hypothesis that the vaginal and gut microbiomes are colonized by a shared community of *Bifidobacterium*, and further emphasize the potential importance of the maternal vaginal microbiome as a source of infant gut microbiota.

Keywords: *Bifidobacterium*, gut, vagina, genome, microbiome, adaptation, specialization, comparative genomics

5.2. Introduction

Bifidobacterium are Gram-positive, non-motile, anaerobic, non-spore forming rod-shaped bacteria. They belong to the *Bifidobacteriaceae* family and are characterized by high genomic G+C content (55-67 mol%) (Biavati and Mattarelli, 2006). Bifidobacteria are common members of the microbiota of the gastrointestinal tract (GIT), and they also colonize the human vagina, oral cavity, and breast milk (Gomez-Gallego *et al.*, 2016; Biavati and Mattarelli, 2006). Beyond the human microbiome, they can be found in sewage, fermented milk products, and the gastrointestinal

tracts of animals including insects (Biavati and Mattarelli, 2012). Although members of the genus inhabit a wide range of habitats, most *Bifidobacterium* spp. are host-specific (Biavati and Mattarelli, 2006).

Bifidobacteria have been the subject of numerous studies due their probiotic potential and health promoting characteristics, such as immune modulation (Fanning et al., 2012; Hart et al., 2004), production of bacteriocins (Martinez et al., 2013) and inhibition of pathogens (Bernet et al., 1993; Arboleya et al., 2011; Fukuda et al., 2011). The precise mechanisms by which bifidobacteria provide these benefits, however, are not fully understood. Bifidobacteria also play an important role as one of the primary colonizers of the neonatal gut, representing 60-91% of fecal bacteria in breast-fed infants (Harmsen et al., 2000; Turroni et al., 2012). This early microbial colonization is an essential step in the modulation of the neonatal immune system (Rautava *et al.*, 2012; Olszak *et al.*, 2012) and may be influenced by mode of delivery (vaginal or C-section) and feeding type (breast milk or formula) (Penders et al., 2006; Dominguez-Bello et al., 2010; Bäckhed et al., 2015).

Studies of the vaginal microbiota using deep sequencing methods have shown that bifidobacteria are the dominant bacteria in the vaginal microbiomes of some reproductive aged women (Freitas and Hill, 2017). Culture-based studies have subsequently confirmed that some vaginal *Bifidobacterium* spp. are able to perform a protective role similar to the beneficial lactobacilli, which includes the production of lactic acid and hydrogen peroxide (Schellenberg et al., 2012; Freitas and Hill, 2017). These features prevent the overgrowth of unwanted bacteria and help to maintain the homeostasis of the vaginal microbiome. Common species detected in the vaginal microbiome, based on sequencing (Chaban et al., 2014; Freitas et al., 2017; Freitas and Hill, 2017) and culture methods (Freitas and Hill, 2017; Verhelst et al., 2005), include *Bifidobacterium breve*, *B. longum*, *B. bifidum*, and *B. dentium*. Based on the recognition that both the gut and vagina harbour bifidobacteria, a large number of studies have been conducted to investigate and demonstrate the influence of maternal microbiota on the neonatal gut microbiome (Matamoros et al., 2013; Bäckhed et al., 2015; Makino et al., 2013; Dominguez-Bello et al., 2010; Gabriel et al., 2017; Gomez-Gallego et al., 2016; Hunt et al., 2011; Milani et al., 2015; Duranti et al., 2017; Pannaraj et al., 2017), but the specific contribution of each microbial community (vaginal, gut and to a smaller extent, milk) in the mother-to-infant bifidobacteria transmission remains in discussion.

Comparison of *Bifidobacterium* spp. genome sequences has revealed a high degree of conservation and synteny across their genomes (Ventura et al., 2007). Nevertheless, phenotypic differences have been described among bifidobacteria, suggesting that niche adaptation has led to some degree of specialization (Ventura et al., 2009a). One indication of these adaptations is the greater percentage of the bifidobacteria genome involved in carbohydrate metabolism in comparison with the genomes of other members of the gut microbiota. Specifically, genome analysis has demonstrated that *B. longum* has the ability to metabolize a variety of complex sugars which gives an ecological advantage in the GIT and evidently reflects its gut adaptation (Schell et al., 2002). It has also been shown that *B. longum* subsp. *infantis* has adaptations for milk utilization (Sela et al., 2008), and that *B. dentium* is adapted for the colonization of the human oral cavity (Ventura et al., 2009b). These observations of niche specialization, however, are mostly based on the examination of a single genome, or reflect the differences among different bifidobacteria species (from the same niche or not) rather than among isolates of the same species from different ecological niches. *B. longum*, for example, possess a large number of genes related to carbohydrate transport and metabolism indicating an adaption to the diversity of nutrients available in the gut (Schell et al., 2002), but since *B. longum* also colonizes the human vagina, it is not clear whether these specializations are broadly present in all *B. longum* or only those strains present in the gut.

The importance of bifidobacteria in adult and neonatal health is evident, although evidence supporting the importance of vertical transmission of maternal microbiota in establishing these populations remains inconclusive. The overlapping occurrence of bifidobacteria species, among other microorganisms, in different body sites is one of the challenges in studying vertical transmission. For example, are vaginal bifidobacteria transferred to the infant gut microbiome, or are they a distinct population adapted to life in the vaginal microbiome? Answers to questions like this are not yet resolved due the lack of studies comparing isolates of the same species from different body sites. Furthermore, while several bifidobacteria adaptations for survival in the GIT and oral cavity have been proposed, no study has addressed possible adaptations to the particular features of the vaginal microbiome. Here, we compared the genomes of gut and vaginal isolates of *B. breve* and *B. longum* to identify evidence of specialization that could indicate if vaginal and gut strains represent two distinct, adapted subpopulations or if they are ubiquitous bacteria able to colonize both gut and vagina. Improved knowledge of bifidobacteria ecology is necessary for a

better understanding of mother-to-infant bifidobacteria transmission, a potentially important determinant of infant health.

5.3. Material and Methods

Bacterial strains

A total of 16 bifidobacteria (7 *Bifidobacterium breve* and 9 *Bifidobacterium longum*) were sequenced in this study. Genome sequences from an additional 27 *Bifidobacterium* spp. were acquired from GenBank for comparative analysis (Table 5.1). All 16 genomes sequenced in this study were from strains originally isolated from human vaginal microbiota as part of previous studies (Freitas and Hill, 2017; Schellenberg et al., 2012). Genomic DNA was isolated from cultures grown in Modified Reinforced Clostridial broth using a modified salting out procedure (Martín-Platero et al., 2007). The integrity of DNA was verified by electrophoresis on 1% agarose gels. Genomic DNA was quantified using Qubit dsDNA BR assay kit (Invitrogen, Burlington, Ontario) and DNA quality was assessed by the A260/A280 ratio using a spectrophotometer.

Genome sequencing and assembly

Libraries were prepared with 1 ng of genomic DNA using the Nextera XT DNA Library Preparation Kit (Illumina Inc., San Diego, CA) according to the manufacturer's instructions. After PCR amplification and clean up, the fragment size distribution of the tagmented DNA was analyzed using the High Sensitivity DNA Analysis Kit on Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). PhiX DNA (15% (v/v)) was added to the pooled indexed libraries prior to loading onto the flow cell. The libraries were sequenced using Reagent Kit V2 (500 cycles) on Illumina Miseq platform (Illumina Inc., San Diego, CA).

Raw sequence reads were trimmed for quality using Trimmomatic (Bolger et al., 2014) with a minimum read length of 40 and quality cut-off of Phred score of 20. To estimate genome coverage and calculate average insert size, reads were mapped on to the reference genome of *B. breve* (Genbank Accession AP012324) or *B. longum* (Genbank Accession NC_015067.1) using Bowtie2 (Langmead and Salzberg, 2012) and the results were converted to BAM format for

viewing in Qualimap v2.2.1 (Okonechnikov et al., 2015). High quality reads were assembled with SOAPdenovo2 (Luo et al., 2012) using the estimated average insert size from Qualimap analysis.

Genome analysis

Genomes were annotated using the NCBI (National Center for Biotechnology Information) Prokaryotic Genome Automatic Annotation Pipeline (PGAAP). CRISPRFinder (http://crispr.i2bc.paris-saclay.fr/) was used to identify CRISPR (clustered regularly interspaced short palindromic repeats) within the genome sequences (Grissa et al., 2008). Annotated genomes were submitted to the Joint Genome Institute (JGI - http://jgi.doe.gov/) for COG (Clusters of Orthologous Groups) category assignment.

Overall genome similarities were assessed by calculating the Average Nucleotide Identity by Mummer (ANIm) and tetranucleotide scores (tetra) within JSpecies (Richter et al., 2016). ANIm values were visualized as heatmap, generated in R. The vegdist() function was used to calculate the Euclidean distance between the ANI divergence values, and hclust() function was used to calculate the complete linkage on the distance matrix, in R.

Pangenome analysis was conducted using GView server with the following settings: genomes were compared to each other using tblastn with expect value cutoff of 1×10^{-10}; CDS 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 80% (Petkau *et al.*, 2010).

Core and accessory genomes were identified using the Nucmer alignment within Spine and the distribution of each accessory genomic element among the genomes was determined with ClustAge (Ozer et al., 2014). A heatmap and neighbour joining tree were created based on the Bray-Curtis distance matrix of the accessory element distribution patterns from the 100 bootstrap replicates.

CSI Phylogeny was used to call SNPs (single-nucleotide polymorphisms) and infer phylogeny based on the concatenated alignment of SNPs. The following settings were used: minimum depth at SNP positions of 10; relative depth at SNP positions of 10; minimum distance between SNPs (prune) of 10; minimum SNP quality of 30; minimum read mapping quality of 25;

minimum Z-score: 1.96. A maximum likelihood tree indicating the whole genome phylogeny was also computed within CSI Phylogeny (Kaas et al., 2014).

5.4. Results and Discussion

General genome features

We performed a comparative genomic analysis of gut and vaginal isolates for two *Bifidobacterium* species commonly found in the human gut and vagina: *B. breve* and *B. longum*. All analysis comparing vaginal and gut strains was performed in parallel for these two species. The general features of all genomes included in this study are listed in Table 5.1. Seventeen *B. breve* genomes were analyzed, seven of which were sequenced in this study. We also analyzed the genome sequences of twenty-six *B. longum* strains, nine of which were sequenced as part of this study. Genome sequence data and annotated assemblies have been deposited in GenBank under BioProject PRJNA387952. The *B. breve* genomes were sequenced to an average of 97-fold coverage ± 47 (range 14-270), and *B. longum* genomes were sequenced to an average of 57-fold coverage ± 34 (range 11-182).

Table 5.1. General features of the bifidobacteria genomes included in this study.

Strain	Ecological origin	Size (Mb)	GC (%)	Total genes	tRNA	CRISPR	N50 / N90	GenBank	Status (n. scaffolds/ contigs)
***B. breve* (n=17)**									
B.b.[#] 30-1	Vagina	2.54	59.8	2346	67	1	125268 / 38442	*	D (32/5)
B.b. 91-1	Vagina	2.24	58.0	2215	50	0	188169 / 56499	*	D (21/1)
B.b. 322-1	Vagina	2.24	58.5	2026	47	1	29270 / 9437	*	D (67/36)
B.b. W20-13	Vagina	2.30	58.3	2113	50	1	43276 / 19775	*	D (49/16)
B.b. W56	Vagina	2.36	57.8	2207	49	2	126027 / 46295	*	D (24/13)
B.b. N6D12	Vagina	2.29	58.5	2097	50	1	119740 / 29243	*	D (30/12)
B.b. 12-4	Vagina	2.26	58.7	2148	48	3	148702 / 48178	*	D (25/7)
B.b. ACS-071-V-Sch8b	Vagina	2.33	58.7	2046	53	4	NA	CP002743	C
B.b. JCM 1192[T]	Infant feces	2.27	58.9	2039	53	0	NA	AP012324	C
B.b. UCC2003	Infant feces	2.42	58.7	2131	54	3	NA	CP000303	C
B.b. JCM 7017	Infant feces	2.29	58.7	1995	54	2	NA	CP006712	C
B.b. JCM 7019	Adult feces	2.36	58.6	2133	56	2	NA	CP006713	C
B.b. NCFB 2258	Infant feces	2.32	58.7	2036	53	2	NA	CP006714	C
B.b. 689b	Infant feces	2.33	58.7	2052	53	0	NA	CP006715	C
B.b. S27	Infant feces	2.29	58.7	2005	53	2	NA	CP006716	C
B.b. CBT BR3	Infant feces	2.43	59.1	2195	54	2	NA	CP010413	C
B.b. LMC520	Infant feces	2.40	59.0	2146	55	1	NA	CP019596	C
***B. longum* (n=26)**									
B.l.[#] 239-2	Vagina	2.28	59.1	2060	50	8	33459 / 10497	*	D (60/56)
B.l. W35-1	Vagina	2.30	60.3	2133	49	1	17831 / 5459	*	D (81/132)
B.l. N2E12	Vagina	2.34	59.1	2114	51	5	80659 / 26287	*	D (43/12)
B.l. N2F05	Vagina	2.33	59.6	2076	55	2	105634 / 41749	*	D (30/13)
B.l. N2G10	Vagina	2.28	59.6	2200	38	0	9284 / 3230	*	D (165/186)
B.l. N3A01	Vagina	2.31	59.7	2056	51	5	82928 / 40450	*	D (33/13)
B.l. N3E01-2	Vagina	2.34	59.6	2167	51	3	202175 / 50971	*	D (23/1)
B.l. N5E04	Vagina	2.46	58.3	2219	45	0	154725 / 47782	*	D (27/3)
B.l. N6D05	Vagina	2.72	60.4	2641	77	3	25630 / 6992	*	D (80/105)
B.l.l. JCM 1217[T]	Infant feces	2.39	60.3	2090	73	1	NA	AP010888	C
B.l.l. JDM301	Gut	2.48	59.8	2156	55	1	NA	CP002010	C
B.l.l. BBMN68	Elderly feces	2.27	59.9	1959	54	3	NA	CP002286	C
B.l.l. KACC 91563	Infant feces	2.40	59.8	2064	56	1	NA	CP002794	C
B.l.l. GT15	Adult feces	2.34	60.0	2021	56	1	NA	CP006741	C
B.l.l. NCIMB 8809	Human feces	2.34	60.1	2037	56	1	NA	CP011964	C
B.l.l. CCUG 30698	Gut	2.46	60.2	2184	72	1	NA	CP011965	C
B.l.l. AH1206	Infant feces	2.42	60.2	2179	60	2	NA	CP016019	C
B.l. NCC2705	Infant feces	2.26	60.1	1799	57	1	NA	AE014295	C
B.l. DJO10A	Adult feces	2.39	60.1	2105	58	2	NA	CP000605	C
B.l. 105-A	Human feces	2.29	60.1	1950	56	2	NA	AP014658	C
B.l. BXY01	Gut	2.48	59.8	2158	55	1	NA	CP008885	C
B.l. BG7	Infant feces	2.46	60.0	2128	57	1	NA	CP010453	C
B.l. 35624	Gut	2.26	60.0	1942	57	2	NA	CP013673	C
B.l.i. JCM 1222[T]	Infant feces	2.83	59.9	2673	77	0	NA	CP001095	C
B.l.i. 157F	Infant feces	2.41	60.1	2147	59	1	NA	AP010890	C
B.l.i. BT1	Infant feces	2.58	59.4	2308	56	1	NA	CP010411	C

[#] B.b. = *B. breve*; B.l. = *B. longum*; B.l.l. = *B. longum* subsp. *longum*; B.l.i. = *B. longum* subsp. *infantis*. * Genome sequenced as part of this study. [T] = type strain. NA = not applicable. C = complete, D = draft.

Currently, *B. longum* encompass three subspecies: -*longum*, -*infantis* and -*suis*, but it has been previously considered as three separate species (*B. longum*, *B. infantis* and *B. suis*) or as a unified species (*B. infantis* and *B. suis* were published as synonyms of *B. longum*) (Mattarelli et al., 2008; Sakata et al., 2002). Considering this controversial taxonomic history of *B. longum*, we opted to include in our analysis publicly complete genomes of all subspecies of *B. longum* of gut origin.

The average genome size of vaginal and gut *B. breve* was 2.32 ± 0.09 Mb and 2.34 ± 0.05 Mb, respectively; where vaginal strains 91-1 and 322-1 were the smallest (2.24 Mb) and the vaginal strain 30-1 was the largest (2.54 Mb). For *B. longum*, the average genome sizes of vaginal and gut isolates were 2.37 ± 0.14 Mb and 2.41 ± 0.13 Mb, respectively; the smallest genomes were represented by gut isolates NCC2705 and 35624 (2.26 Mb) and the largest genome was from JCM 1222^{T} (2.83 Mb). All genomes analyzed had high GC content, a known characteristic of the genus *Bifidobacterium* and previously reported as 55-67 mol% (Biavati and Mattarelli, 2012). Vaginal and gut *B. breve* have 58.5% and 58.8% genomic GC content, respectively; and vaginal and gut *B. longum* have 59.5% and 60% of GC content, respectively. There were no differences in genome size and GC content between gut and vaginal isolates, either for *B. breve* or *B. longum* (t-test, all $p>0.05$) (Figure 5.1).

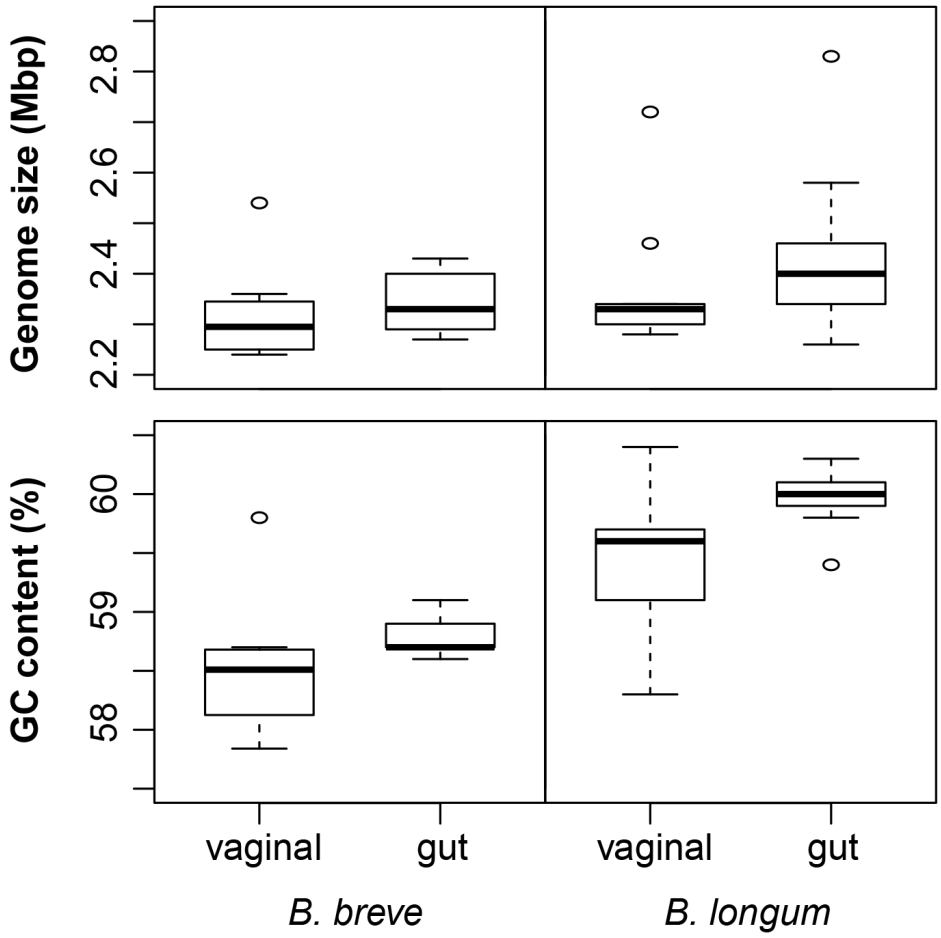

Figure 5.1. Genome features.

Genome size and GC content of *B. breve* and *B. longum* of gut and vaginal origin.

The genomes of *B. breve* and *B. longum* contained an average of 2113 and 2137 genes, respectively, which is within the range of number of predicted genes previously reported for *Bifidobacterium* spp. (1369-2564 genes) (Sun et al., 2015). Also, *B. breve* and *B. longum* contained an average of 1.6 (range 0-4) and 1.8 (range 0-8) CRISPR, respectively. There was no difference in the number of genes and CRISPR between gut and vaginal isolates of either *B. breve* or *B. longum* (t-test, all $p>0.05$). CRISPR have been implicated in chromosomal rearrangement, modulation of expression of neighbouring genes, target for DNA binding proteins, and DNA repair (Horvath and Barrangou, 2010). It has been shown to act as the defense mechanism in bacteria against phages and plasmids by providing adaptive immunity (Barrangou et al., 2007). Notably, vaginal *B. longum* 239-2 and N2E12 had a total of 8 and 5 CRISPR, respectively, which suggest these strains had an active CRISPR immune system against potentially damaging foreign DNA. Previous studies have shown that CRISPR systems are frequent and diverse in the genus *Bifidobacterium* and differences in the frequency of CRISPR-Cas systems (CRISPR and CRISPR-associated proteins) within species are an indicative that CRISPR distribution is strain-dependent (Briner et al., 2015; Milani et al., 2014).

The overall sequence similarity between vaginal and gut strains was assessed based on average nucleotide identity (ANI). A hierarchal clustering based on the distance matrix of ANI divergence values was computed and visualized as a heatmap (Figure 5.2). All ANI values between isolates of the same species were >95%, consistent with their identification as members of the same species (Richter and Rossello-Mora, 2009). The genomes of vaginal *B. breve* did not cluster separately from the gut isolates (Figure 5.2A). Similarly, vaginal *B. longum* did not form a separate cluster from the gut isolates (Figure 5.2B). Therefore, vaginal and gut isolates could not be distinguished based on their overall nucleotide identity, either for *B. breve* or *B. longum*. The relatedness of isolates was also assessed by calculating the tetranucleotide scores (tetra), which led to similar results as ANI (data not shown).

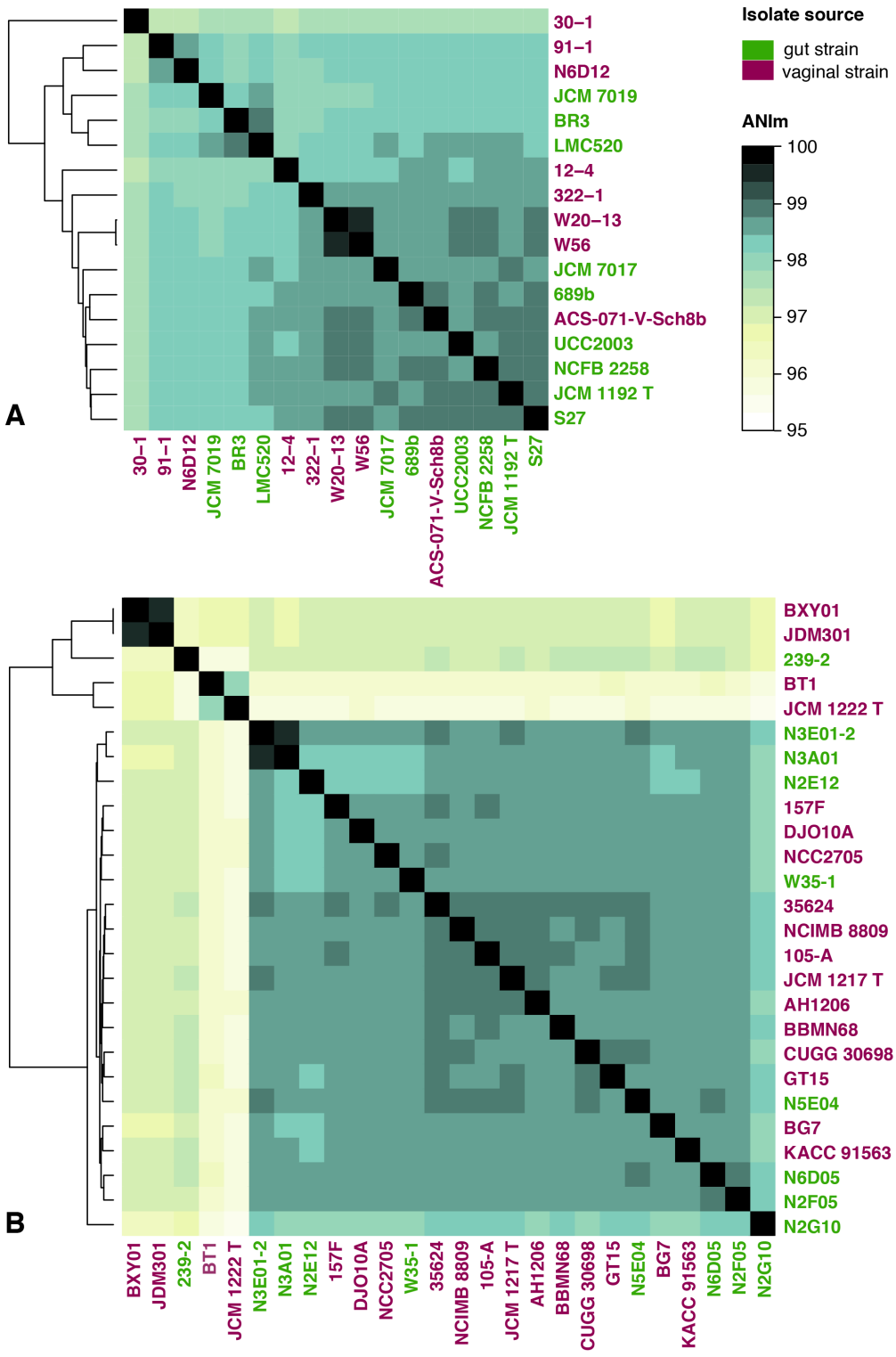

Figure 5.2. Average Nucleotide Identity (ANI).

Heatmap of ANI values between genomes of gut and vaginal origin. (A) *B. breve*; (B) *B. longum*. T = type strain.

Pangenome, core and accessory genome

The pangenome is defined as the entire gene set of all isolates, including genes present in all isolates (core genome) and genes present in one or some isolates (accessory genome). The pangenome analysis revealed that genomes of *B. breve* and *B. longum* did not have major deletions or insertions related to the source of the isolates (Figure 5.3). In other words, at this overall level, there was no region within the genome specifically associated with gut or vaginal isolates.

Interestingly, a few vaginal and gut *B. breve* isolates had several coding sequences in common that were mostly absent in other isolates (Figure 5.3A). The gut isolate JCM 1192^T and the vaginal isolates 30-1 and N6D12 shared genes encoding several hypothetical proteins, a type IA DNA topoisomerase, a plasmid mobilization relaxosome protein MobC, a mobilization protein, a PsfB family protein, a conjugal transfer protein TraG, a glycosidase hydrolase, a transfer complex protein, and a Prgl family protein, all located within the 1300-1350 kbp region of the calculated pangenome (Figure 5.3A). In the same region, several hypothetical proteins, a cell wall anchor protein and a ParA family protein were also exclusively present in the gut strain JCM 1192^T and vaginal strain 30-1. These two isolates also shared an integrase, a type I restriction modification protein subunit S, a SAM-dependent DNA methyltransferase, and a type I desoxyribonuclease HsdR, all of which were located within the 1650-1700 kbp region (Figure 5.3A).

Analysis of *B. longum* pangenome also revealed several sequence features shared by the vaginal and gut isolates (Figure 5.3B). In particular, the gut isolates JCM 1217^T, CCUG 30698, KACC 91563, NCIMB 8809, 35624 and JCM 1222^T, and the vaginal isolates 239-2 and N5E04 shared several coding sequences within the region 1550-1555 Kbp (Figure 5.3B); among the shared sequences were numerous hypothetical proteins, a DUF262 domain containing protein, a cell division protein FtsZ, a DNA helicase UvrD, a ATP-dependent nuclease subunit B, and a DNA polymerase III subunit alpha. Notably, vaginal isolate 239-2 was missing several coding sequences in comparison with other isolates, which is consistent with its small size (2.28 Mb). Moreover, gut isolates JCM 1222^T and BT1 had very similar genome patterns, where presence/absence of several sequences was in common between them.

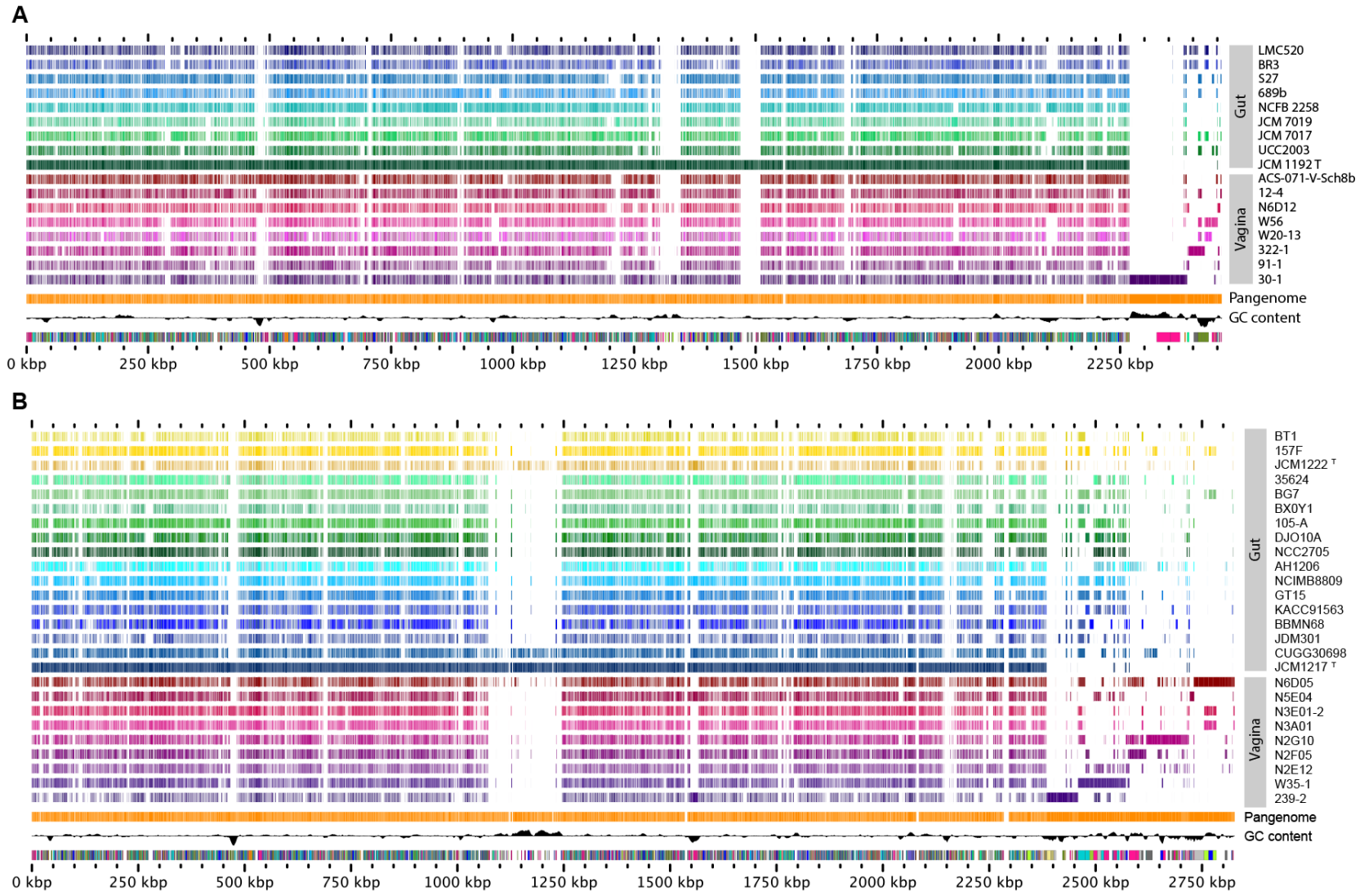

Figure 5.3. Pangenome.

(A) *B. breve*; (B) *B. longum*. Coloured horizontal bars represent different strains. Vertical grey blocks indicate origin of the isolate. T = type strain.

We also investigated *B. breve* and *B. longum* core genomes, which represent the DNA coding sequences (CDS) that are present in all isolates of a species. For *B. breve* isolates, the core genome accounted for an average of 69.8% of the genome (1.6 Mb), for both vaginal and gut isolates, ranging from 63.3% (strain 30-1) to 73.2% (strain 91-1) for the vaginal isolates, and from 66% (strain BR3) to 70.8% (strain JCM 1192^T) for the gut isolates. For *B. longum*, the core genome accounted for an average of 55.7% (1.33 Mb) of the genome (vaginal isolates: 56.4%; gut isolates: 55.3%), ranging from 49.1% (strain N6D05) to 58.3% (strain 239-2) for vaginal isolates, and from 46.2% (strain JCM 1222^T) to 58.9% (strain NCC2705) for the gut isolates. No significant differences in the proportional representation of the core genome of gut and vaginal isolates were detected either for *B. breve* or *B. longum* (t-test, all $p>0.05$).

The accessory genome was defined in terms of presence or absence of CDS within the genome; a Bray-Curtis dissimilarity matrix was constructed based on the distribution of these accessory elements. The genomes of *B. breve* and *B. longum* did not cluster based on the origin of the strains (Figure 5.4), indicating that the presence/absence of accessory genes did not correlate according to the isolates source. Notably, the relationship of the isolates based on the accessory genome is consistent with observations in the pangenome analysis. Strain 30-1 and JCM 1192^T, for example, clustered together based on the accessory elements (Figure 5.3A) and had several extra genes in common in the pangenome analysis. Also, vaginal strain 239-2 did not cluster with any other isolate based on the accessory genome analysis (Figure 5.3B), which is consistent with the fact that several coding sequences were missing in this strain in the pangenome analysis making it more different from the other isolates.

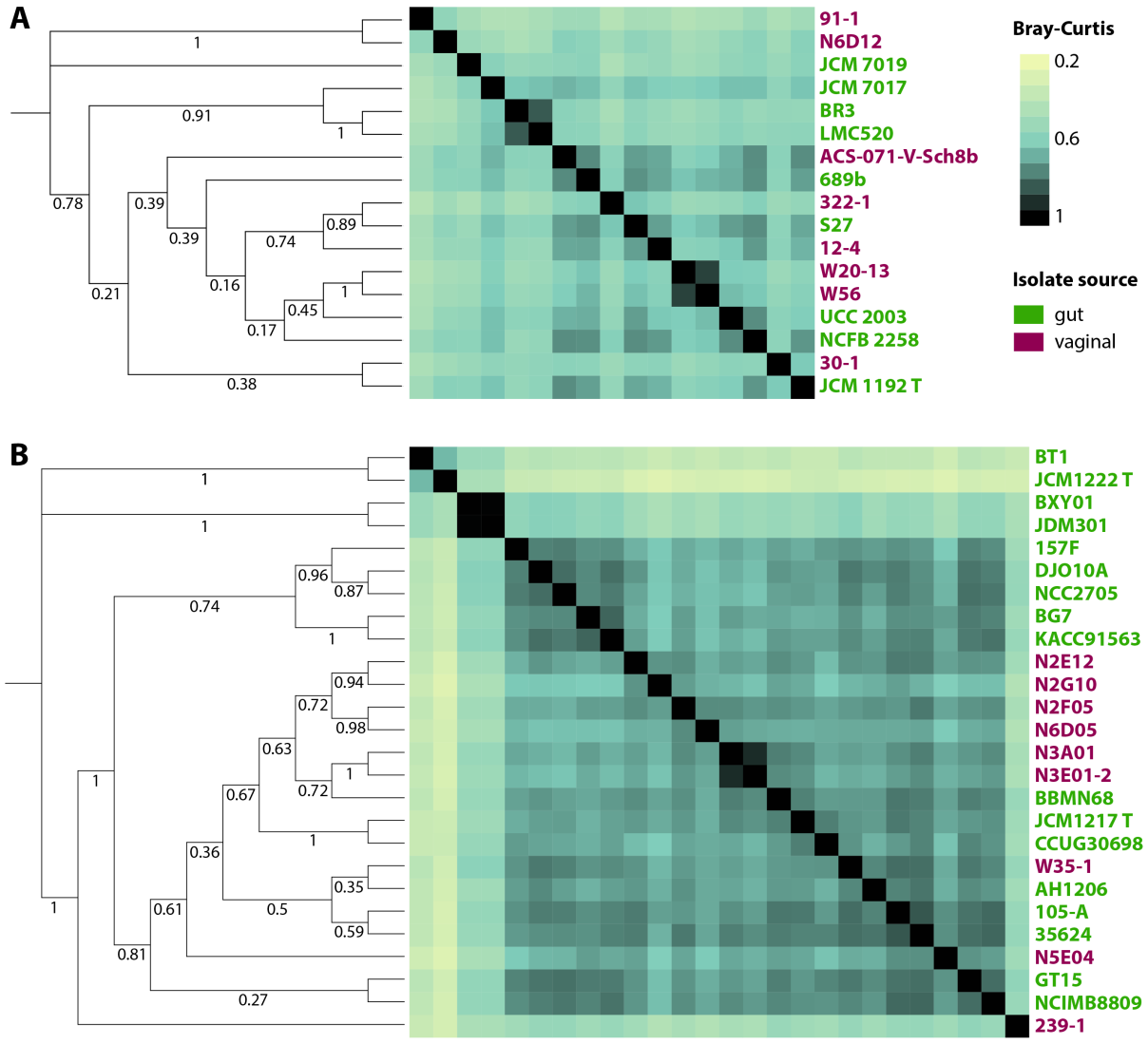

Figure 5.4. Accessory genome.

Heatmap and neighbour joining tree based on Bray-Curtis distance matrix of the accessory genome elements. (A) *B. breve*; (B) *B. longum*. T = type strain.

Whole genome phylogeny

Whole genome phylogeny of isolates was inferred based on the concatenated alignment of SNPs, and a maximum likelihood circular tree was created for phylogeny visualization (Figure 5.5). The comparison of all genomes revealed a total of 1,616,032 and 1,350,284 SNPs in all *B. breve* and *B. longum* strains, respectively. Phylogenetic analysis indicated that vaginal *B. breve* isolates did not cluster separately from the gut isolates (Figure 5.5A). Similarly, the genomes of vaginal *B. longum* did not form a separate cluster from the genomes of gut strains (Figure 5.5B).

COG distribution

Predicted proteins were functionally categorized based on COG (Cluster of Orthologous Group) assignment and the proportions in each category were compared between vaginal and gut isolates (Figure 5.6). Most sequences in *B. breve* and *B. longum* were assigned to the 'carbohydrate transport and metabolism' category, followed by 'amino acid transport and metabolism' and 'translation, ribosomal structure and biogenesis' categories (Figure 5.6A-B). Most importantly, vaginal and gut strains did not differ in terms of COG category distribution. Considering that carbohydrates are less abundant in quantity and variety in the vagina relative to the gut, we had anticipated that if vaginal isolates have genome adaptations, it would be mostly related to genes involved in metabolism. Previous studies have demonstrated that bifidobacteria isolated from different hosts have different COG distribution and are enriched with specific gene sets according to their niche. *B. boum* and *B. asteroides*, species usually isolated from bees, have a group of genes encoding oxidases involved in respiration metabolism, an adaptation for an environment richer in oxygen (Sun et al., 2015). Another study has shown that *B. dentium* isolated from the oral cavity has more COG assigned to the toxin-defense category relative to other enteric bifidobacteria (Ventura et al., 2009b). In this study, however, there was no indication that vaginal bifidobacteria, *B. breve* or *B. longum*, possess different proportions of genes across different COG functions.

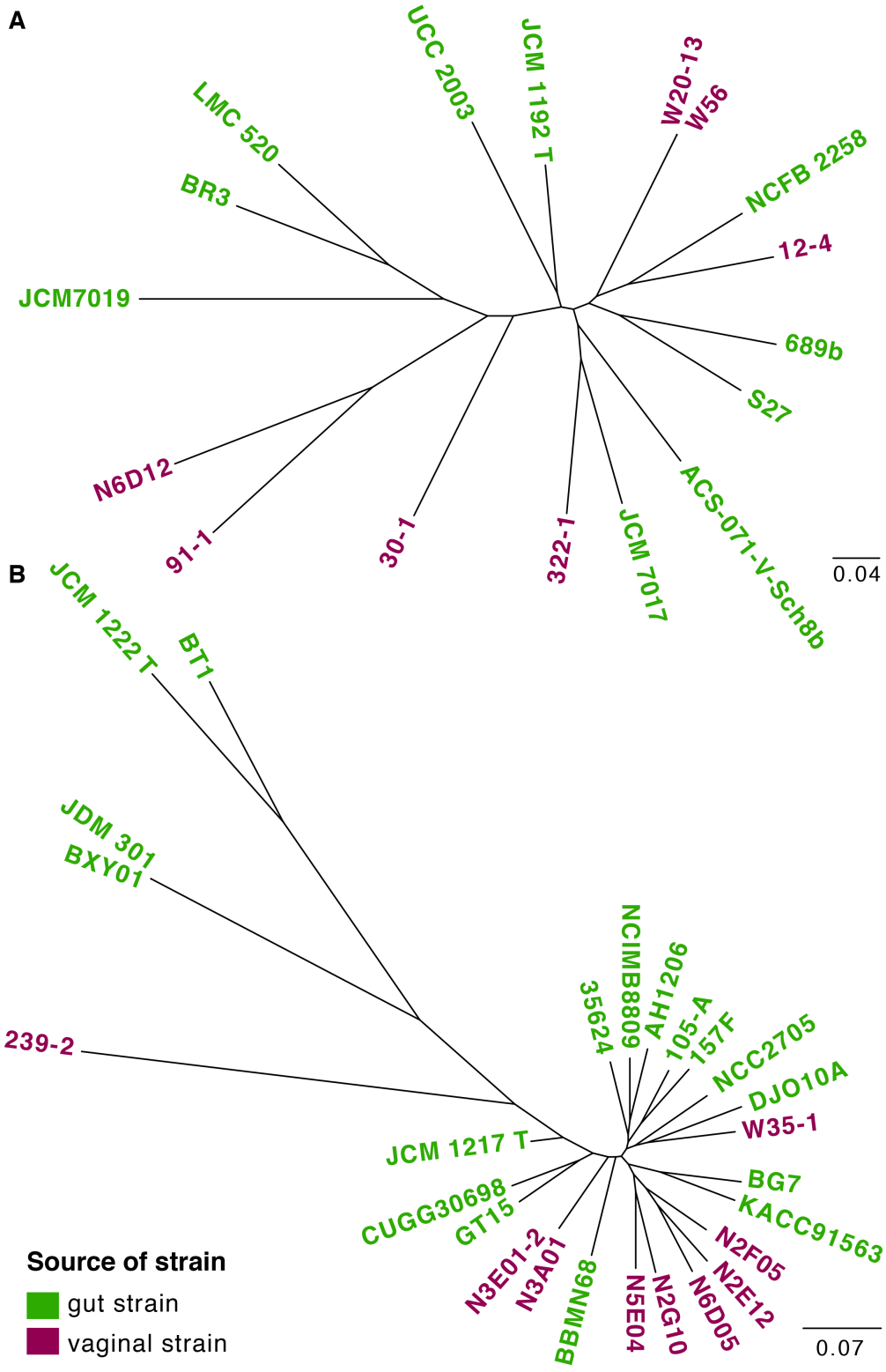

Figure 5.5. Whole genome phylogeny.

A maximum likelihood tree based on SNPs. (A) *B. breve*; (B) *B. longum*. T = type strain.

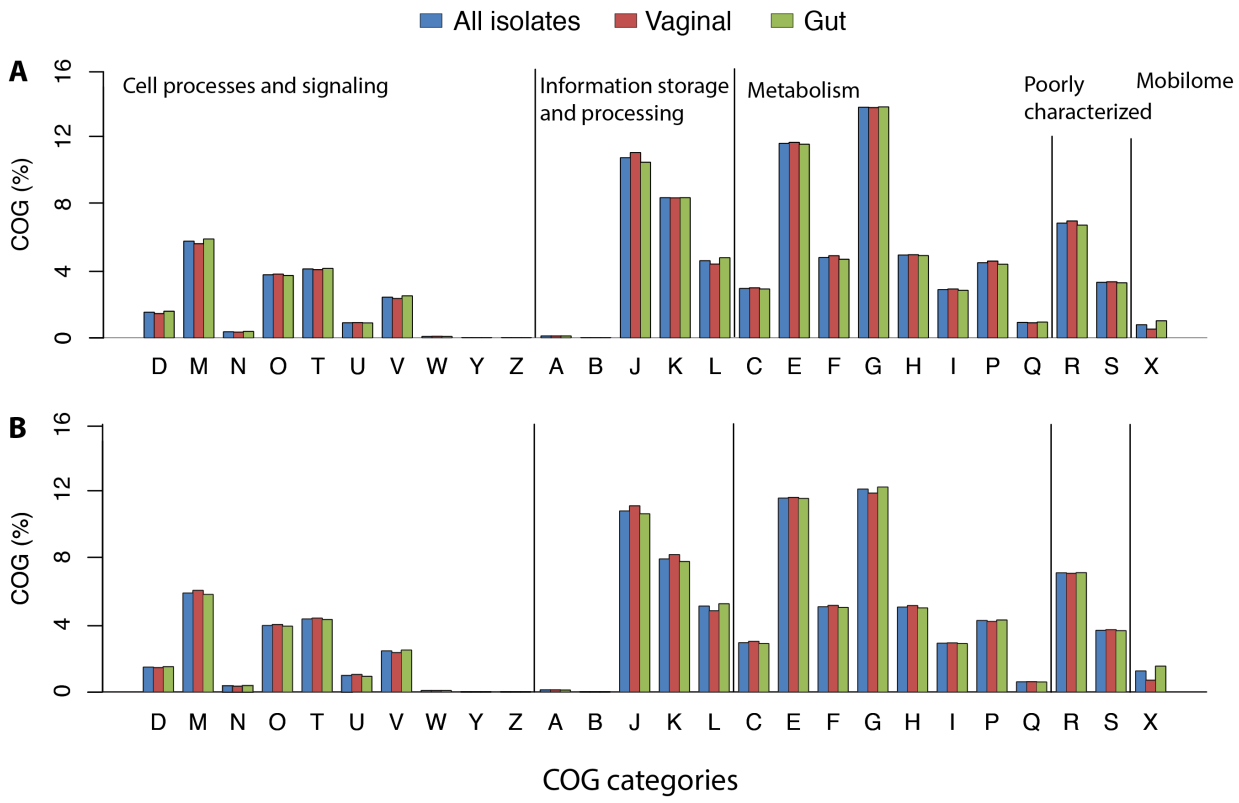

Figure 5.6. Clusters of Orthologous Groups (COG).

COG function distribution among genomes of (A) *B. breve* and (B) *B. longum*. COG classification: [D] Cell cycle control, cell division, chromosome partitioning; [M] Cell wall/membrane/envelope biogenesis; [N] Cell motility; [O] Post-translational modification, protein turnover, and chaperones; [T] Signal transduction mechanisms; [U] Intracellular trafficking, secretion, and vesicular transport; [V] Defense mechanisms; [W] Extracellular structures; [Y] Nuclear structure; [Z] Cytoskeleton; [A] RNA processing and modification; [B] Chromatin structure and dynamics; [J] Translation, ribosomal structure and biogenesis; [K] Transcription; [L] Replication, recombination and repair; [C] Energy production and conversion; [E] Amino acid transport and metabolism; [F] Nucleotide transport and metabolism; [G] Carbohydrate transport and metabolism; [H] Coenzyme transport and metabolism; [I] Lipid transport and metabolism; [P] Inorganic ion transport and metabolism; [Q] Secondary metabolites biosynthesis, transport, and catabolism; [R] General function prediction only; [S] Function unknown; [X] Mobilome: prophages, transposons.

COG categories were also investigated in terms of number of shared and "unique" COG among isolates. *B. breve* isolates were represented by a set of 1016 COG, 489 of which were present in all strains (vaginal and gut origin). A total of 30 COG were exclusively present in vaginal isolates, i.e., they were absent in gut isolates and present in at least one vaginal isolate. Notably, the prevalence of these 30 "unique" COG among vaginal isolates was low, ranging from 12.5% (1/8) to 37.5% (3/8), suggesting it is unlikely that these rare COG represent a biologically significant vaginal adaptation. Similarly, a total of 34 COG were exclusively present in gut isolates, but their prevalence was also mostly low, ranging from 11.1% (1/9) to 33.3% (3/9), with only one exception: COG1396 (Transcriptional regulator, contains XRE-family HTH domain) that was present in 55.5% (5/9) of the gut isolates. Two additional COG showed substantial differences in distribution between vaginal and gut *B. breve* isolates. Glucan phosphorylase (COG0058) was present in only 12.5% (1/8) of vaginal isolates against 88.9% (8/9) of isolates of gut origin. On the other hand, the predicted ABC-type sugar transport system (permease component) (COG4158) was present in 87% (7/8) and 22.2% (2/9) of vaginal and gut isolates, respectively.

For *B. longum*, a total of 1128 COG were identified, 451 of which were present in all strains (vaginal and gut). Although 26 COG were exclusively present in vaginal isolates, most of them were found in only one or two vaginal isolates, with one exception: COG3695 (Alkylated DNA nucleotide flippase Atl1), which was present in 55.5% (5/9) of vaginal isolates. On the other hand, a total of 113 COG were exclusively associated with isolates of gut origin, although their prevalence was also low, ranging from 5.8% (1/17) to 29.4% (5/17). The only exception was COG1672 (Predicted ATPase), which was found in 76.5% (13/17) of gut strains. Additionally, 5 COG were more frequently found in gut isolates than vaginal isolates: COG0481 (Translation elongation factor EF-4, membrane-bound GTPase), COG0802 (tRNA A37 threonylcarbamoyladenosine biosynthesis protein TsaE), COG1225 (Peroxiredoxin), COG0159 (Tryptophan synthase alpha chain), and COG0732 (Restriction endonuclease S subunit). Four COG were more prevalent in vaginal isolates: COG1327 (Transcriptional regulator NrdR, contains Zn-riB.b.on and ATP-cone domains), COG0328 (Ribonuclease HI), COG1983 (Phage shock protein PspC (stress-responsive transcriptional regulator)), and COG0759 (Membrane-anchored protein YidD, putatitve component of membrane protein insertase Oxa1/YidC/SpoIIIJ). Although there were a few differences in COG distribution between vaginal and gut isolates, differences

were not systematically concentrated within function categories, and were insufficient to distinguish isolates from the two body sites.

5.5. Conclusions

In this study, we investigated the genomes of *B. breve* and *B. longum* from two different body sites of significant importance in neonatal health: gut and vagina. In all analyses, gut and vaginal strains of *B. breve* or *B. longum* were not distinguishable from each other based on their genomic content. Consistent with these observations, it has been previously demonstrated that several vaginal and gut *Bifidobacterium* spp. did not differ based on phenotypic characteristics related to their carbohydrate fermentation patterns, lactic acid production and tolerance to low pH (Freitas and Hill, 2017). Our results support the hypothesis that vaginal and gut isolates represent the same bacterial population with similar genetic repertoires that allow them to efficiently colonize both body sites. The genomes included in our study are from isolates recovered from individuals of different ages around the world over many years. While this provides an opportunity to look at these species very broadly, an obvious complementary study would be a comparison of gut and vaginal isolates from individual women, or paired samples from women and their babies. Both vaginal and gut microbiota are thought to contribute to mother-infant transmission of bifidobacteria, and recognition that vaginal and gut isolates represent the same population is an important step for future studies in maternal-neonatal health.

Acknowledgments

The authors would like to thank Champika Fernando for the help with genome sequencing and Sarah Vancuren for assistance in genome assembly.

CHAPTER 6. General discussion and conclusions

6.1. Summary and limitations of these works

Pregnancy affects the vaginal microbiota by increasing relative abundance of lactobacilli, decreasing microbiota richness and diversity, and decreasing prevalence of Mollicutes

The vaginal microbiome has been the subject of many studies over the years. These studies have clearly demonstrated that the vaginal microbial community influences women's reproductive health. A disturbed microbiota is associated with increased susceptibility to sexually transmitted infections and negative reproductive outcomes. Despite advances in technology, the causes of disturbed microbiota are still poorly understood. In addition, most studies have focused in the microbiota of reproductive age, non-pregnant women. Pregnancy is associated with a variety of physiological events that may, or may not drive changes in the composition of the vaginal microbiota. Therefore, a better understanding of the vaginal microbiome in low risk pregnancy is necessary before further investigating the microbiota as a tool for prediction of adverse pregnancy outcomes.

In the study described in Chapter 2, a baseline description of the vaginal microbiota in pregnant women at low risk for preterm birth was established, based on sequencing of the cpn60 universal target. cpn60 provides higher resolution than 16S rRNA variable regions (Links et al., 2012) and allowed the detection of novel CST that had been previously described based on the resolution of distinct subgroups of *G. vaginalis*, a hallmark bacteria in disturbed vaginal microbial communities (Albert et al., 2015). The relatively large sample size and the cpn60 approach are considered strengths that differentiate this work from other studies. Our results demonstrated that pregnancy was associated with decreased microbiota richness and diversity, decreased Mollicutes and *Ureaplasma* prevalence, and an increased relative abundance of lactobacilli and total bacterial load. These pregnancy-associated changes were hypothesized to be related to increased levels of estrogen that have been previously shown to increase deposition of glycogen in vaginal epithelial cells (BMJ, 1943; Cruickshank and Sharman, 1934). Glycogen is the main carbohydrate utilized by vaginal lactobacilli for the production of lactic acid (Cruickshank, 1934), which may provide a competitive advantage for vaginal lactobacilli and explain the greater dominance of these bacteria, and consequently lower prevalence of other bacteria such as Mollicutes, and overall lower richness and diversity in the vaginal microbiota of pregnant women in comparison to non-pregnant women.

One of the potential limitations of this study is the single time point collection since it does not provide information about possible changes in microbial composition as a function of gestational age. The results of this study are only applied to pregnant women early in gestation and it is possible that the changes in microbial composition reported here might be different for other gestation periods. Another limitation is the lack of information about the weight of vaginal swabs. The results indicated that microbiomes of pregnant women contained more total bacteria than the microbiomes of non-pregnant women. This might have a biological significance (increased glycogen may provide a more nutritious environment that supports growth of more bacteria) or might be simply the result of more material been collected from pregnant women in comparison with non-pregnant women. This question unfortunately remains unsolved since swabs were not weighed before DNA extraction steps.

Increased microbiota richness and diversity is associated with preterm birth

Imbalances in the vaginal microbiota have long been associated with negative reproductive outcomes such as spontaneous preterm birth, but the mechanisms underlying the association between a disturbed microbiota and preterm remain poorly understood. Preterm birth is a multifactorial condition in which an intrauterine infection ascending from the vagina is thought to be an important contributor to the onset of preterm labour. Nevertheless, no single bacterium or group of bacteria have yet been causally linked to prematurity, which demonstrates the complexity of this condition.

In the study described in Chapter 3, the microbiota of women who had term and preterm deliveries (gestational age at delivery <37 weeks) were compared based on their microbial profiles generated by sequencing of the cpn60 universal target. The overall goal was to identify traits in the microbiota that could be further investigated as predictors of preterm birth. Results of this study demonstrated that the microbiota of women who experienced preterm birth had higher richness and diversity, and higher Mollicutes prevalence when compared to those of women who delivered at term. At an overall level, no specific microbial profile pattern was associated with preterm birth, which unfortunately does not contribute directly to the development of a diagnostic tool based on the microbiota for the prediction of preterm birth. Nevertheless, the increased richness and diversity (in the preterm cohort) might be a remnant characteristic of a previously disturbed microbiota, or an early indication of a microbiota shift from a healthy to a disturbed status. Either

way, this increased microbiota richness/diversity might be an indication that the bacterial community overwhelmed placental and amniotic defences and permitted ascending bacterial contamination of the chorioamnion, which probably led to preterm birth.

One limitation of this study is related to single time point sample collection. Other studies have shown that changes in the microbiota based on microbial profile patterns are very transient (DiGiulio et al., 2015), which makes detection the non-*Lactobacillus*-dominated profiles more difficult if samples are not collected frequently. A longitudinal study would have provided the opportunity to capture transient, disturbed microbiomes that could be associated with preterm birth. Another limitation is the lack of extreme preterm cases. Most women in this study were considered late preterm (average gestational week was 34) and although this question cannot be addressed, it is possible that preterm driven by an ascending infection would be more evident in extreme preterm cases.

Bifidobacteria can dominate the vaginal microbiota and have the potential to be as beneficial as lactobacilli in promoting a healthy vaginal microbiome

The healthy vaginal microbiome is currently understood to be dominated by lactobacilli, and an overgrowth of any other species is often associated with unhealthy conditions. As a result of application of DNA sequencing based methods for microbiome characterization, an increasing appreciation of atypical microbiomes in healthy women has emerged, such as those dominated by bifidobacteria. Bifidobacteria are well known for their health promoting characteristics in the gastro-intestinal tract, but relatively little is known about vaginal bifidobacteria and their potential to promote a healthy vaginal environment.

In the first part of this study (Chapter 4), quantitative PCR was used to confirm the relative abundance of *Bifidobacterium* in vaginal microbiomes of several women, previously determined by deep sequencing. Primer bias in PCR amplification for microbial profiling might over/underestimate specific groups of bacteria. Here, the results indicated that bifidobacteria are the dominant members of the vaginal microbiota in some women and that previous results based on deep sequencing were in fact valid. Thereafter, vaginal bifidobacteria were isolated and phenotypically characterized in the context of features believed to promote reproductive health. Most isolates were identified as *B. breve* or *B. longum*, based on identity of the cpn60 gene. Results demonstrated that some bifidobacteria produced levels of lactic acid similar to the beneficial

lactobacilli. While only 32% of isolates produced hydrogen peroxide, most vaginal bifidobacteria were able to tolerate high levels of lactic acid (100 mM) and low pH (4.5 or 3.9), conditions typical of a healthy vaginal microbiome. These results collectively demonstrated that *Bifidobacterium* spp. have the potential to be as protective as lactobacilli according to the current understanding of a healthy vaginal microbiome. The results also contributed to improving the definition of a healthy microbiome and may help to guide clinicians and researchers to better assess a healthy vaginal microbiome, avoiding unnecessary interventions.

A major limitation of this study is the low number of isolates per species. A total of 8 different bifidobacteria species were isolated and characterized, which was a good representation of the bifidobacteria species commonly detected in vaginal microbiome. However, only *B. breve* (n=15) and *B. longum* (n=11) had a significant number of isolates that could be analyzed statistically throughout the experiments, while results derived from remaining isolates were largely descriptive. Having more isolates per species would have provided more information about significant differences in the phenotypic characteristics among different vaginal bifidobacteria species.

Vaginal and gut microbiomes are colonized by a shared community of *Bifidobacterium*

Bifidobacteria colonize the human gastrointestinal tract, vagina, oral cavity and breast milk. They influence human physiology through gut health-promoting effects and play an important role as primary colonizers of the newborn gut, which contributes for the maturation of the neonatal immune system. Bifidobacteria are acquired both vertically (from the mother) and horizontally (from other humans or the environment) during and after birth. Nevertheless, the contribution of each maternal microbiome remains in discussion since the mere presence of common bacterial species does not necessarily indicate vertical transmission, and studies using high-resolution methods for strain characterization are still limited.

B. breve and *B. longum* are both commonly detected in the human vaginal and gut microbiome, but it is not yet clear if they represent distinct, niche-adapted subpopulations. In this study (Chapter 5), the genomes of *B. breve* and *B. longum* from vaginal and gut microbiomes were investigated. The goal was to assess if vaginal and gut bifidobacteria represent the same population, which could have implications in understanding vertical transmission. Vaginal and gut isolates of either *B. breve* or *B. longum* did not differ in size, GC content, number of genes and

CRISPR. In addition, vaginal and gut isolates did not cluster separately based on average nucleotide identity or whole genome phylogeny analysis. They also had a similar COG category distribution with no evidence of differential adaptation to either body site. Therefore, the results of this study supported the hypothesis that the vaginal and gut microbiomes are colonized by a shared community of *Bifidobacterium*, and emphasized the importance of the maternal vaginal microbiome as a source of infant gut microbiota.

Limitations of this study relate to the lack of gut and vaginal isolates from the same women or paired samples from women and their babies. This could provide more information about the contribution of each body site to bifidobacteria vertical transmission.

6.2. Discussion of future prospects

The vaginal microbiome has been studied over the past years mostly by applying culture dependent techniques. Advances in the development of high-throughput, next-generation sequencing have revolutionized characterization of human vaginal microbiota. These culture-independent methods provide identification of bacteria at low taxonomic levels, such as genus and species, and allow detection of bacterial species that are difficult to culture. This represents a great advance for vaginal microbiome characterization studies in comparison with low-resolution methods like the Nugent score, which is based on Gram-stains of vaginal smears. Despite these technological advances, the causes of microbiome dysbiosis and its potential association with preterm birth are not completely understood.

In Chapters 2 and 3 of this thesis, DNA sequencing based methods were used to investigate the vaginal microbiota of different cohorts. Results of these studies provided important information about the composition of the vaginal microbiomes of a Canadian population of healthy pregnant women, and pregnant women who experienced preterm birth. Nevertheless, it is known that the incidence of microbiome dysbiosis and preterm birth varies among women of different ethnicities suggesting that host factors may also play a role in influencing the microbiota and reproductive outcomes. Additional studies of the vaginal microbiomes of women from different geographical locations and ethnicities are essential to better understand the relationship among host factors, microbiota and reproductive outcomes. Gene-environment studies have also suggested that preterm birth following a bacterial infection might be host-response related. Some women may

hyper-respond to a uterine infection by releasing large amounts of inflammatory cytokines, which could potentially trigger premature labour (Hernandez-Guerrero et al., 2003; Crider et al., 2005). The carriage of specific maternal alleles of inflammation-related genes (e.g., TNF-2, tumour necrosis factor alpha gene) in combination with a disturbed vaginal microbiota has also been proposed as a synergistic mechanism for an higher risk of preterm birth in comparison with each risk factor alone (Macones et al., 2004; Gómez et al., 2010). Therefore, assessing microbiota composition is only one part of the equation in trying to understand what defines healthy and high-risk microbiomes for preterm birth. Future studies investigating host and environmental factors in association with the microbiota could improve understanding of the causes of preterm birth and other negative reproductive outcomes.

The definition of a healthy vaginal microbiota is also a limiting factor to successfully diagnosing women with a disturbed microbial community. The current definition is simplistic and out-dated, but evidence supporting a new and better definition is still limited. In Chapter 4 of this thesis, it was confirmed that bifidobacteria are the dominant members of some vaginal microbiomes and that bifidobacteria have ecological and phenotypic characteristics similar to the beneficial vaginal lactobacilli, such as high production of lactic acid and tolerance to low pH. In addition of acidifying the vaginal environment, lactobacilli have also been demonstrated to (i) be able to adhere to cervical and vaginal epithelial cells forming a biological barrier against colonization by pathogenic bacteria (Coudeyras et al., 2008), (ii) inhibit growth of urogenital pathogens (Atassi *et al.*, 2006), and (iii) block pathogens adherence through process related to competition for vaginal epithelium receptors, displacement and co-aggregation (Osset et al., 2001; Castro et al., 2013; Kaewnopparat et al., 2013; Boris et al., 1998), (Kaewnopparat et al., 2013). Although less investigated, these characteristics represent important aspects of lactobacilli colonization that contribute to promoting a healthy vaginal microbiome. In vaginal bifidobacteria, these characteristics have not yet been assessed. In vitro studies investigating the ability of bifidobacteria to adhere to vaginal epithelial cells and to inhibit growth and adherence of urogenital pathogens are essential to confirm that bifidobacteria, like lactobacilli, contribute to vaginal homeostasis.

The importance of vaginal bifidobacteria has often been overlooked, however, gut bifidobacteria have been extensively studied due their health promoting characteristics and probiotic potential in the gastro-intestinal tract. In newborns, bifidobacteria are one of the most

abundant members of the gut microbiota and contribute to early immune system maturation. The results reported in Chapter 5 of this thesis indicated that gut and vaginal microbiomes share a population of bifidobacteria. The comparison of genome sequences of vaginal and gut isolates indicated that vaginal and gut microbiomes are colonized by a shared community of *Bifidobacterium* spp., which emphasizes the contribution of the vaginal microbiota as a source of bacteria for the neonatal gut. The contribution of maternal microbiota is undeniable, and many studies have demonstrated that the gut of neonates and their mothers harbour similar (bifido)bacteria species, which does not necessarily mean that they were vertically transmitted. This gap in knowledge related to the specific contribution of each maternal microbiota in vertical transmission might be clarified with further studies using high-resolution methods to compare the bifidobacteria species, at strain level, from the same women (from gut and vagina) or from paired samples of women and their babies.

REFERENCES

Aagaard KM, Ma J, Antony KM, Ganu R, Petrosino J, Versalovic J. (2014). The placenta harbors a unique microbiome. *Sci Transl Med* **6**: 1–11.

Aagaard KM, Riehle K, Ma J, Segata N, Mistretta TA, Coarfa C, *et al.* (2012). A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. *PLoS One* **7**: e36466.

AbdelRahman YM, Belland RJ. (2005). The chlamydial developmental cycle. *FEMS Microbiol Rev* **29**: 949–959.

Agrawal V, Hirsch E. (2012). Intrauterine infection and preterm labor. *Semin Fetal Neonatal Med* **17**: 12–19.

Alakomi HL, Skyttä E, Saarela M, Mattila-Sandholm T, Latva-Kala K, Helander IM. (2000). Lactic acid permeabilizes gram-negative bacteria by disrupting the outer membrane. *Appl Environ Microbiol* **66**: 2001–2005.

Alauzet C, Mory F, Teyssier C, Hallage H, Carlier JP, Grollier G, *et al.* (2010). Metronidazole resistance in *Prevotella* spp. and description of a new *nim* gene in *Prevotella baroniae*. *Antimicrob Agents Chemother* **54**: 60–64.

Albert AYK, Chaban B, Wagner EC, Schellenberg JJ, Links MG, van Schalkwyk J, *et al.* (2015). A study of the vaginal microbiome in healthy canadian women utilizing cpn60-based molecular profiling reveals distinct *Gardnerella* subgroup community state types. *PLoS One* **10**: e0135620.

Aldunate M, Tyssen D, Johnson A, Zakir T, Sonza S, Moench TR, *et al.* (2013). Vaginal concentrations of lactic acid potently inactivate HIV. *J Antimicrob Chemother* **68**: 2015–2025.

Allsworth JE, Peipert JF. (2011). Severity of bacterial vaginosis and the risk of sexually transmitted infection. *Am J Obstet Gynecol* **205**: 113.e1-113.e6.

Amsel R, Totten PA, Spiegel C, Chen K, Eschenbach DA, Holmes KK. (1983). Nonspecific vaginitis. Diagnostic criteria and microbial and epidemiologic associations. *Am J Med* **74**: 14–22.

Andrews WW, Goldenberg RL, Hauth JC, Cliver SP, Copper R, Conner M. (2006a). Interconceptional antibiotics to prevent spontaneous preterm birth: a randomized clinical trial. *Am J Obstet Gynecol* **194**: 617–623.

Andrews WW, Goldenberg RL, Mercer B, Iams J, Meis P, Moawad A, *et al.* (2000). The preterm prediction study: association of second-trimester genitourinary chlamydia infection with subsequent spontaneous preterm birth. *Am J Obstet Gynecol* **183**: 662–668.

Andrews WW, Klebanoff MA, Thom EA, Hauth JC, Carey JC, Meis PJ, *et al.* (2006b). Midpregnancy genitourinary tract infection with *Chlamydia trachomatis*: association with subsequent preterm delivery in women with bacterial vaginosis and *Trichomonas vaginalis*. *Am J Obstet Gynecol* **194**: 493–500.

Arboleya S, Ruas-Madiedo P, Margolles A, Solís G, Salminen S, de los Reyes-Gavilán CG, *et al.* (2011). Characterization and in vitro properties of potentially probiotic *Bifidobacterium* strains isolated from breast-milk. *Int J Food Microbiol* **149**: 28–36.

Aroutcheva AA, Simoes JA, Behbakht K, Faro S. (2001). *Gardnerella vaginalis* isolated from patients with bacterial vaginosis and from patients with healthy vaginal ecosystems. *Clin Infect Dis* **33**: 1022–1027.

Arrieta MC, Stiemsma LT, Amenyogbe N, Brown EM, Finlay B. (2014). The intestinal microbiome in early life: health and disease. *Front Immunol* **5**: 1–18.

Asnicar F, Manara S, Zolfo M, Truong DT, Scholz M, Armanini F, *et al.* (2017). Studying vertical microbiome transmission from mothers to infants by strain-level metagenomic profiling. *mSystems* **2**: 1–13.

Atassi F, Brassart D, Grob P, Graf F, Servin AL. (2006). *Lactobacillus* strains isolated from the vaginal microbiota of health women inhibit *Prevotella bivia* and *Gardnerella vaginalis* in coculture and cell culture. *FEMS Immunol Med Microbiol* **48**: 424–432.

Bäckhed F, Ding H, Wang T, Hooper L V, Koh GY, Nagy A, *et al.* (2004). The gut microbiota as an environmental factor that regulates fat storage. *Proc Natl Acad Sci U S A* **101**: 15718–15723.

Bäckhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, *et al.* (2015). Dynamics and stabilization of the human gut microbiome during the first year of life. *Cell Host Microbe* **17**: 690–703.

Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, *et al.* (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**: 1709–1712.

Bautista CT, Wurapa EK, Sateren WB, Morris SM, Hollingsworth BP, Sanchez JL. (2016). Association of bacterial vaginosis with chlamydia and gonorrhea among women in the U.S. army. *Am J Prev Med* **52**: 632–639.

Behrman RE, Butler AS. (2007). Preterm birth: causes, consequences, and prevention. National Academies Press: Washington, D.C.

Beigi RH, Austin MN, Meyn LA, Krohn MA, Hillier SL. (2004). Antimicrobial resistance associated with the treatment of bacterial vaginosis. *Am J Obstet Gynecol* **191**: 1124–1129.

Benjamini Y, Hochberg Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc* **57**: 289–300.

Bernet MF, Brassart D, Neeser JR, Servin AL. (1993). Adhesion of human bifidobacterial strains to cultured human intestinal epithelial cells and inhibition of enteropathogen-cell interactions. *Appl Environ Microbiol* **59**: 4121–4128.

Biavati B, Mattarelli P. (2012). Genus *Bifidobacterium*. In: Goodfellow M, Kämpfer P, Busse H-J, Trujillo ME, Suzuki K, Ludwig W, et al. (eds). *Bergey's Manual® of Systematic Bacteriology*

2nd, volume 5. Springer: New York, pp 171–205.

Biavati B, Mattarelli P. (2006). The Family Bifidobacteriaceae. In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E (eds). *The Prokaryotes: Volume 3: Archaea. Bacteria: Firmicutes, Actinomycetes*. Springer: New York, pp 322–382.

Blencowe H, Cousens S, Oestergaard MZ, Chou D, Moller AB, Narwal R, *et al.* (2012). National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet* **379**: 2162–2172.

BMJ. (1943). Oestrogens and vaginal glycogen. *Br Med J* **2**: 753.

Bobetsis YA, Barros SP, Offenbacher S. (2006). Exploring the relationship between periodontal disease and pregnancy complications. *J Am Dent Assoc* **137**: S7–S13.

Bolger AM, Lohse M, Usadel B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Bollopragada S, Youssef R, Jordan F, Greer I, Norman J, Nelson S. (2009). Term labor is associated with a core inflammatory response in human fetal membranes, myometrium, and cervix. *Am J Obstet Gynecol* **200**: 104.e1-104.e11.

Bondici VF, Lawrence JR, Khan NH, Hill JE, Yergeau E, Wolfaardt GM, *et al.* (2013). Microbial communities in low permeability, high pH uranium mine tailings: characterization and potential effects. *J Appl Microbiol* **114**: 1671–1686.

Borgdorff H, Gautam R, Armstrong S, Xia D, Ndayisaba G, van Teijlingen N, *et al.* (2015). Cervicovaginal microbiome dysbiosis is associated with proteome changes related to alterations of the cervicovaginal mucosal barrier. *Sex Transm Infect* **91**: 621–633.

Borgdorff H, Tsivtsivadze E, Verhelst R, Marzorati M, Jurriaans S, Ndayisaba GF, *et al.* (2014). *Lactobacillus*-dominated cervicovaginal microbiota associated with reduced HIV/STI prevalence and genital HIV viral load in African women. *ISME J* **8**: 1781–1793.

Boris S, Barbés C. (2000). Role played by lactobacilli in controlling the population of vaginal pathogens. *Microbes Infect* **2**: 543–546.

Boris S, Suárez JE, Vázquez F, Barbés C. (1998). Adherence of human vaginal lactobacilli to vaginal epithelial cells and interaction with uropathogens. *Infect Immun* **66**: 1985–1989.

Boskey E, Cone RA, Whaley K, Moench TR. (2001). Origins of vaginal acidity: high D/L lactate ratio is consistent with bacteria being the primary source. *Hum Reprod* **16**: 1809–1813.

Boskey E, Telsch K, Whaley K, Moench TR, Cone RA. (1999). Acid production by vaginal flora in vitro is consistent with the rate and extent of vaginal acidification. *Infect Immun* **67**: 5170–5175.

Bottacini F, Ventura M, Sinderen D, Motherway M. (2014). Diversity, ecology and intestinal function of bifidobacteria. *Microb Cell Fact* **13**: 1–15.

Bradshaw CS, Morton AN, Hocking J, Garland SM, Morris MB, Moss LM, *et al.* (2006). High recurrence rates of bacterial vaginosis over the course of 12 months after oral metronidazole therapy and factors associated with recurrence. *J Infect Dis* **193**: 1478–1486.

Breshears LM, Edwards VL, Ravel J, Peterson ML. (2015). *Lactobacillus crispatus* inhibits growth of *Gardnerella vaginalis* and *Neisseria gonorrhoeae* on a porcine vaginal mucosa model. *BMC Microbiol* **15**: 1–12.

Bretelle F, Rozenberg P, Pascal A, Favre R, Bohec C, Loundou A, *et al.* (2015). High *Atopobium vaginae* and *Gardnerella vaginalis* vaginal loads are associated with preterm birth. *Clin Infect Dis* **60**: 860–867.

Briner AE, Lugli GA, Milani C, Duranti S, Turroni F, Gueimonde M, *et al.* (2015). Occurrence and diversity of CRISPR-Cas systems in the genus *Bifidobacterium*. *PLoS One* **10**: e0133661.

Briselden ANNM, Moncla BJ, Stevens CE, Hillier SL. (1992). Sialidases (neuraminidases) in bacterial vaginosis and bacterial vaginosis-associated microflora. *J Clin Microbiol* **30**: 663–666.

Brooks JP, Edwards DJ, Harwich MD, Rivera MC, Fettweis JM, Serrano MG, *et al.* (2015). The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiol* **15**: 1–14.

Brotman RM. (2011). Vaginal microbiome and sexually transmitted infections: an epidemiologic perspective. *J Clin Invest* **121**: 4610–4617.

Brotman RM, He X, Gajer P, Fadrosh D, Sharma E, Mongodin EF, *et al.* (2014a). Association between cigarette smoking and the vaginal microbiota: a pilot study. *BMC Infect Dis* **14**: 1–11.

Brotman RM, Ravel J, Bavoil PM, Gravitt PE, Ghanem KG. (2014b). Microbiome, sex hormones, and immune responses in the reproductive tract: challenges for vaccine development against sexually transmitted infections. *Vaccine* **32**: 1543–1552.

Brousseau R, Hill JE, Prefontaine G, Goh S, Harel J, Hemmingsen SM. (2001). *Streptococcus suis* serotypes characterized by analysis of chaperonin 60 gene sequences. *Appl Environ Microbiol* **67**: 4828–4833.

Brunham RC, Rey-Ladino J. (2005). Immunology of *Chlamydia* infection: implications for a *Chlamydia trachomatis* vaccine. *Nat Rev Immunol* **5**: 149–161.

Budd W, Bostwick DG, Woody J, Hunt C. (2016). Antimicrobial resistance genes and modelling of treatment failure in bacterial vaginosis: clinical study of 289 symptomatic women. *J Med Microbiol* **65**: 377–386.

Burton JP, Cadieux PA, Reid G. (2003a). Improved understanding of the bacterial vaginal microbiota of women before and after probiotic instillation. *Appl Environ Microbiol* **69**: 97–101.

Burton JP, Dixon JL, Reid G. (2003b). Detection of *Bifidobacterium* species and *Gardnerella vaginalis* in the vagina using PCR and denaturing gradient gel electrophoresis. *Int J Gynecol*

Obstet **81**: 61–63.

Buvé A, Jespers V, Crucitti T, Fichorova RN. (2014). The vaginal microbiota and susceptibility to HIV. *AIDS* **28**: 2333–2344.

Callahan BJ, DiGiulio DB, Goltsman DSA, Sun CL, Costello EK, Jeganathan P, *et al.* (2017). Replication and refinement of a vaginal microbial signature of preterm birth in two racially distinct cohorts of US women. *Proc Natl Acad Sci U S A* **114**: 9966–9971.

Capoccia R, Greub G, Baud D. (2013). *Ureaplasma urealyticum, Mycoplasma hominis* and adverse pregnancy outcomes. *Curr Opin Infect Dis* **26**: 231–240.

Caporaso JG, Kuczynski J, Stombaugh JI, Bittinger K, Bushman FD, Costello EK, *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Carey J, Blackwelder W, Nugent R, Matteson M, Rao A, Eschenbach D, *et al.* (1992). Antepartum cultures for *Ureaplasma urealyticum* are not useful in predicting pregnancy outcome. *Int J Gynecol Obstet* **37**: 153–153.

Castro J, Henriques A, Machado A, Henriques M, Jefferson KK, Cerca N. (2013). Reciprocal interference between *Lactobacillus* spp. and *Gardnerella vaginalis* on initial adherence to epithelial cells. *Int J Med Sci* **10**: 1193–1198.

Cauci S, Culhane JF. (2011). High sialidase levels increase preterm birth risk among women who are bacterial vaginosis–positive in early gestation. *Am J Obstet Gynecol* **204**: 142.e1-142.e9.

Cauci S, Monte R, Ropele M, Missero C, Not T, Quadrifoglio F, *et al.* (1993). Pore-forming and haemolytic properties of the *Gardnerella vaginalis* cytolysin. *Mol Microbiol* **9**: 1143–1155.

CDC - Centers for Disease Control and Prevention. (2015a). Sexually transmitted disease surveillance 2014. U.S. Department of Health and Human Services: Atlanta.

CDC - Centers for Disease Control and Prevention. (2015b). Sexually transmitted diseases treatment guidelines, 2015. In: Vol. 64. *MMWR Recommendations and Reports*. U.S. Department of Health and Human Services: Atlanta.

Chaban B, Albert AYK, Links MG, Gardy J, Tang P, Hill JE. (2013). Characterization of the upper respiratory tract microbiomes of patients with pandemic H1N1 influenza. *PLoS One* **8**: e69559.

Chaban B, Links MG, Jayaprakash TP, Wagner EC, Bourque DK, Lohn Z, *et al.* (2014). Characterization of the vaginal microbiota of healthy Canadian women through the menstrual cycle. *Microbiome* **2**: 1–12.

Chaplin A V, Efimov BA, Smeianov V V, Kafarskaia LI, Pikina AP, Shkoporov AN. (2015). Intraspecies genomic diversity and long-term persistence of *Bifidobacterium longum*. *PLoS One* **10**: e0135658.

Chen C, Song X, Wei W, Zhong H, Dai J, Lan Z, *et al.* (2017). The microbiota continuum along

the female reproductive tract and its relation to uterine-related diseases. *Nat Commun* **8**: 1–11.

Chenoll E, Casinos B, Bataller E, Astals P, Echevarría J, Iglesias JR, *et al.* (2011). Novel probiotic *Bifidobacterium bifidum* CECT 7366 strain active against the pathogenic bacterium *Helicobacter pylori*. *Appl Environ Microbiol* **77**: 1335–1343.

Cherpes TL, Hillier SL, Meyn LA, Busch JL, Krohn MA. (2008). A delicate balance: risk factors for acquisition of bacterial vaginosis include sexual activity, absence of hydrogen peroxide-producing lactobacilli, black race, and positive herpes simplex virus type 2 serology. *Sex Transm Dis* **35**: 78–83.

Cherpes TL, Meyn LA, Krohn MA, Lurie JG, Hillier SL. (2003). Association between acquisition of herpes simplex virus type 2 in women and bacterial vaginosis. *Clin Infect Dis* **37**: 319–325.

Cho I, Blaser MJ. (2012). The human microbiome: at the interface of health and disease. *Nat Rev Genet* **13**: 260–270.

Chrétien FC. (1978). Ultrastructure and variations of human cervical mucus during pregnancy and the menopause. *Acta Obstet Gynecol Scand* **57**: 337–348.

Claman P, Toye B, Peeling RW, Jessamine P, Belcher J. (1995). Serologic evidence of *Chlamydia trachomatis* infection and risk of preterm birth. *C Can Med Assoc J* **153**: 259–262.

Clifton VL, Stark MJ, Osei-Kumah A, Hodyl N a. (2012). Review: The feto-placental unit, pregnancy pathology and impact on long term maternal health. *Placenta* **33**: S37–S41.

CLSI. (2014). Performance standards for antimicrobial susceptibility testing; 24th Informational Supplement. CLSI document M100-S24. Clinical and Laboratory Standards Institute: Wayne, PA.

Conde-Agudelo A, Rosas-Bermúdez A, Kafury-Goeta AC. (2006). Birth spacing and risk of adverse perinatal outcomes. *JAMA* **295**: 1809–1823.

Conti C, Malacrino C, Mastromarino P. (2009). Inhibition of herpes simplex virus type 2 by vaginal lactobacilli. *J Physiol Pharmacol* **60 Suppl 6**: 19–26.

Costa MO, Chaban B, Harding JCS, Hill JE. (2014). Characterization of the fecal microbiota of pigs before and after inoculation with *'Brachyspira hampsonii'*. *PLoS One* **9**: e106399.

Coudeyras S, Jugie G, Vermerie M, Forestier C. (2008). Adhesion of human probiotic *Lactobacillus rhamnosus* to cervical and vaginal cells and interaction with vaginosis-associated pathogens. *Infect Dis Obstet Gynecol* **2008**: 1–5.

Cox C, Watt AP, Mckenna JP, Coyle P V. (2016). *Mycoplasma hominis* and *Gardnerella vaginalis* display a significant synergistic relationship in bacterial vaginosis. *Eur J Clin Microbiol Infect Dis* **35**: 481–487.

Crider KS, Whitehead N, Buus RM. (2005). Genetic variation associated with preterm birth: A HuGE review. *Genet Med* **7**: 593–604.

Crozier SR, Robinson SM, Borland SE, Godfrey KM, Cooper C, Inskip HM. (2009). Do women change their health behaviours in pregnancy? Findings from the Southampton women's survey. *Paediatr Perinat Epidemiol* **23**: 446–453.

Cruickshank R. (1934). The conversion of the glycogen of the vagina into lactic acid. *J Pathol Bacteriol* **39**: 213–219.

Cruickshank R, Sharman A. (1934). The biology of the vagina in the human subject. II The bacterial flora and secretion of the vagina at various age-periods and their relations to glycogen in the vaginal epitehlial. *BJOG An Int J Obstet Gynaecol* **41**: 190–207.

Csángó PA. (1982). First international conference on vaginosis: nonspecific vaginitis. *Scand J Infect Dis* **14**: 1–126.

Dehnart J. (1957). Untersuchungen úber die gram positive Stuhlflora des Brust- milchkinder. *Zentrabl Bakteriol Parasitenkd Infekt Hyg Abt I Orig R A* **169**: 66–79.

Desai AR, Links MG, Collins SA, Mansfield GS, Drew MD, Van Kessel AG, *et al.* (2012). Effects of plant-based diets on the distal gut microbiome of rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* **350**–**353**: 134–142.

Devillard E, Burton JP, Hammond J, Lam D, Reid G. (2004). Novel insight into the vaginal microflora in postmenopausal women under hormone replacement therapy as analyzed by PCR-denaturing gradient gel electrophoresis. *Eur J Obstet Gynecol Reprod Biol* **117**: 76–81.

DiGiulio DB, Callahan BJ, Mcmurdie PJ, Costello EK, Lyell DJ, Robaczewska A, *et al.* (2015). Temporal and spatial variation of the human microbiota during pregnancy. *Proc Natl Acad Sci U S A* **112**: 11060–11065.

DiGiulio DB, Romero R, Amogan HP, Kusanovic JP, Bik EM, Gotsch F, *et al.* (2008). Microbial prevalence, diversity and abundance in amniotic fluid during preterm labor: a molecular and culture-based investigation. *PLoS One* **3**: e3056.

DiGiulio DB, Romero R, Kusanovic JP, Gómez R, Kim CJ, Seok KS, *et al.* (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. *Am J Reprod Immunol* **64**: 38–57.

Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, *et al.* (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A* **107**: 11971–11975.

Donders GG, Van Bulck B, Caudron J, Londers L, Vereecken A, Spitz B. (2000). Relationship of bacterial vaginosis and mycoplasmas to the risk of spontaneous abortion. *Am J Obstet Gynecol* **183**: 431–437.

Donders GG, Van Calsteren K, Bellen G, Reybrouck R, Van Den Bosch T, Riphagen I, *et al.* (2009). Predictive value for preterm birth of abnormal vaginal flora, bacterial vaginosis and

aerobic vaginitis during the first trimester of pregnancy. *BJOG An Int J Obstet Gynaecol* **116**: 1315–1324.

Donlan RM, Costerton JW. (2002). Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin Microbiol Rev* **15**: 167–193.

Dudley DJ, Collmer D, Mitchell MD, Trautman MS. (1996). Inflammatory cytokine mRNA in human gestational tissues: implications for term and preterm labor. *J Soc Gynecol Investig* **3**: 328–335.

Dumonceaux TJ, Hill JE, Pelletier CP, Paice MG, Van Kessel AG, Hemmingsen SM. (2006). Molecular characterization of microbial communities in Canadian pulp and paper activated sludge and quantification of a novel *Thiothrix eikelboomii*-like bulking filament. *Can J Microbiol* **52**: 494–500.

Duranti S, Lugli GA, Mancabelli L, Armanini F, Turroni F, James K, *et al.* (2017). Maternal inheritance of bifidobacterial communities and bifidophages in infants through vertical transmission. *Microbiome* **5**: 1–13.

Duranti S, Turroni F, Lugli GA, Milani C, Viappiani A, Mangifesta M, *et al.* (2014). Genomic characterization and transcriptional studies of the starch-utilizing strain *Bifidobacterium adolescentis* 22L. *Appl Environ Microbiol* **80**: 6080–6090.

Easmon C, Hay P, Ison C. (1992). Bacterial vaginosis: a diagnostic approach. *Genitourin Med* **68**: 134–138.

Edwards RK, Ferguson RJ, Reyes L, Brown M, Theriaque DW, Duff P. (2006). Assessing the relationship between preterm delivery and various microorganisms recovered from the lower genital tract. *J Matern Neonatal Med* **19**: 357–363.

Egan M, O'Connell Motherway M, Kilcoyne M, Kane M, Joshi L, Ventura M, *et al.* (2014). Cross-feeding by *Bifidobacterium breve* UCC2003 during co-cultivation with *Bifidobacterium bifidum* PRL2010 in a mucin-based medium. *BMC Microbiol* **14**: 1–14.

Endres DM, Schindelin JE. (2003). A new metric for probability distributions. *IEEE Trans Inf Theory* **49**: 1858–1860.

Eschenbach DA, Davick PR, Williams BL, Klebanoff SJ, Young-Smith K, Critchlow CM, *et al.* (1989). Prevalence of hydrogen peroxide-producing *Lactobacillus* species in normal women and women with bacterial vaginosis. *J Clin Microbiol* **27**: 251–256.

Eschenbach DA, Thwin SS, Patton DL, Hooton TM, Stapleton AE, Agnew K, *et al.* (2000). Influence of the normal menstrual cycle on vaginal tissue, discharge, and microflora. *Clin Infect Dis* **30**: 901–907.

Fanaro S, Chierici R, Guerrini P, Vigi V. (2003). Intestinal microflora in early infancy: composition and development. *Acta Paediatr Supplemment* **92**: 48–55.

Fanning S, Hall LJ, Cronin M, Zomer A, MacSharry J, Goulding D, *et al.* (2012). Bifidobacterial surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection. *Proc Natl Acad Sci U S A* **109**: 2108–2113.

FAO, WHO. (2002). Guidelines for the evaluation of probiotics in food. Report of a Joint FAO/WHO Working Group on Drafting Guidelines for the Evaluation of Probiotics in Food. Ontario, Canada.

Farage MA, Miller KW, Sobel JD. (2010). Dynamics of the vaginal ecosystem - Hormonal influences. *Infect Dis Res Treat* **3**: 1–15.

Fashemi B, Delaney ML, Onderdonk AB, Fichorova RN. (2013). Effects of feminine hygiene products on the vaginal mucosal biome. *Microb Ecol Heal Dis* **24**: 1–6.

Fayol-Messaoudi D, Berger CN, Coconnier-Polter M-H, Lievin-Le Moal V, Servin AL. (2005). pH-, lactic acid-, and non-lactic acid-dependent activities of probiotic lactobacilli against *Salmonella enterica* serovar *Typhimurium*. *Appl Environ Microbiol* **71**: 6008–6013.

Felsenstein J. (1993). PHYLIP (Phylogeny Inference Package) version 3.5c, distributed by the author, Department of Genetics, University of Washington, Seattle, USA. http://evolution.genetics.washington.edu/phylip.html.

Fernandes AD, Macklaim JM, Linn TG, Reid G, Gloor GB. (2013). ANOVA-like differential expression (ALDEx) analysis for mixed population RNA-Seq. *PLoS One* **8**: e67019.

Fettweis JM, Brooks JP, Serrano MG, Sheth NU, Girerd PH, Edwards DJ, *et al.* (2014). Differences in vaginal microbiome in African American women versus women of European ancestry. *Microbiology* **160**: 2272–2282.

Fichorova RN, Yamamoto HS, Delaney ML, Onderdonk AB, Doncel GF. (2011). Novel vaginal microflora colonization model providing new insight into microbicide mechanism of action. *MBio* **2**: 1–10.

Foxman B, Wen A, Srinivasan U, Goldberg D, Marrs CF, Owen J, *et al.* (2014). Mycoplasma, bacterial vaginosis-associated bacteria BVAB3, race, and risk of preterm birth in a high-risk cohort. *Am J Obstet Gynecol* **210**: 226.e1-226.e7.

Fraser AM, Brockert JE, Ward RH. (1995). Association of young maternal age with adverse reproductive outcomes. *N Engl J Med* **332**: 1113–1118.

Fredricks DN, Fiedler TL, Thomas KK, Mitchell CM, Marrazzo JM. (2009). Changes in vaginal bacterial concentrations with intravaginal metronidazole therapy for bacterial vaginosis as assessed by quantitative PCR. *J Clin Microbiol* **47**: 721–726.

Fredricks DN, Fiedler TL, Thomas KK, Oakley BB, Marrazzo JM. (2007). Targeted PCR for detection of vaginal bacteria associated with bacterial vaginosis. *J Clin Microbiol* **45**: 3270–3276.

Freitas AC, Chaban B, Bocking A, Rocco M, Yang S, Hill JE, *et al.* (2017). The vaginal

microbiome of pregnant women is less rich and diverse, with lower prevalence of Mollicutes, compared to non-pregnant women. *Sci Rep* **7**: 1–16.

Freitas AC, Hill JE. (2017). Quantification, isolation and characterization of *Bifidobacterium* from the vaginal microbiomes of reproductive aged women. *Anaerobe* **47**: 145–156.

Fukuda S, Toh H, Hase K, Oshima K, Nakanishi Y, Yoshimura K, *et al.* (2011). Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* **469**: 543–547.

Gabriel I, Olejek A, Stencel-Gabriel K, Wielgoś M. (2017). The influence of maternal vaginal flora on the intestinal colonization in newborns and 3-month-old infants. *J Matern Neonatal Med* **7058**: 1–6.

Gajer P, Brotman RM, Bai G, Sakamoto J, Schütte UME, Zhong X, *et al.* (2012). Temporal dynamics of the human vaginal microbiota. *Sci Transl Med* **4**: 1–12.

Gallo MF, Macaluso M, Warner L, Fleenor ME, Hook EW, Brill I, *et al.* (2012). Bacterial vaginosis, gonorrhea, and chlamydial Infection among women attending a sexually transmitted disease clinic: a longitudinal analysis of possible causal links. *Ann Epidemiol* **22**: 213–220.

Gao W, Weng J, Gao Y, Chen X. (2013). Comparison of the vaginal microbiota diversity of women with and without human papillomavirus infection: a cross-sectional study. *BMC Infect Dis* **13**: 1–10.

Gardella C, Riley DE, Hitti J, Agnew K, Krieger JN, Eschenbach DA. (2004). Identification and sequencing of bacterial rDNAs in culture-negative amniotic fluid from women in premature labor. *Am J Perinatol* **21**: 319–323.

Gardner HL, Dukes CD. (1955). Haemophilus vaginalis vaginitis. *Am J Obstet Gynecol* **69**: 962–976.

Gareau MG, Sherman PM, Walker WA. (2010). Probiotics and the gut microbiota in intestinal health and disease. *Nat Rev Gastroenterol Hepatol* **7**: 503–514.

Gerber S, Vial Y, Hohlfeld P, Witkin SS. (2003). Detection of *Ureaplasma urealyticum* in second-trimester amniotic fluid by polymerase chain reaction correlates with subsequent preterm labor and delivery. *J Infect Dis* **187**: 518–521.

Gibbs RS. (1993). Chorioamnionitis and bacterial vaginosis. *Am J Obstet Gynecol* **169**: 460–462.

Goldenberg RL, Culhane JF, Iams JD, Romero R. (2008). Epidemiology and causes of preterm birth. *Lancet* **371**: 75–84.

Goldenberg RL, Hauth JC, Andrews WW. (2000). Intrauterine infection and preterm delivery. *N Engl J Med* **342**: 1500–1507.

Goldenberg RL, Klebanoff MA, Nugent RP, Krohn MA, Hillier SL, Andrews WW. (1996). Bacterial colonization of the vagina during pregnancy in four ethnic groups. *Am J Obstet Gynecol*

174: 1618–1621.

Gomez-Gallego C, Garcia-Mantrana I, Salminen S, Carmen Collado M. (2016). The human milk microbiome and factors in fluencing its composition and activity. *Semin Fetal Neonatal Med* **21**: 400–405.

Gómez LM, Sammel MD, Appleby DH, Elovitz MA, Baldwin DA, Jeffcoat MK, *et al.* (2010). Evidence of a gene-environment interaction that predisposes to spontaneous preterm birth: a role for asymptomatic bacterial vaginosis and DNA variants in genes that control the inflammatory response. *Am J Obstet Gynecol* **202**: 386.e1-386.e6.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, *et al.* (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **29**: 644–652.

Graver MA, Wade JJ. (2011). The role of acidification in the inhibition of *Neisseria gonorrhoeae* by vaginal lactobacilli during anaerobic growth. *Ann Clin Microbiol Antimicrob* **10**: 1–5.

Greenwood JR, Pickett MJ. (1980). Transfer of *Haemophilus vaginalis* Gardner and Dukes to a new genus, *Gardnerella: G. vaginalis* (Gardner and Dukes) comb. nov. *Int J Syst Bacteriol* **30**: 170–178.

Gregoire AT, Kandil O, Ledger WJ. (1971). The glycogen content of human vaginal epithelial tissue. *Fertil Steril* **22**: 64–68.

Grissa I, Vergnaud G, Pourcel C. (2008). CRISPRcompar: a website to compare clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* **36**: W52–W57.

Hall V, Copsey SD. (2015). *Propionibacterium, Lactobacillus, Actinomyces*, and other non-sporing forming anaerobic Gram-positive rods. In: Jorgensen JH, Pfaller MA (eds). *Manual of Clinical Microbiology*. American Society of Microbiology: Washington, D.C., pp 920–939.

Hamark B, Forssman L. (1991). Postabortal endometritis in chlamydia-negative women - Association with preoperative clinical signs of infection. *Gynecol Obstet Invest* **31**: 102–105.

Han Z, Mulla S, Beyene J, Liao G, McDonald SD. (2011). Maternal underweight and the risk of preterm birth and low birth weight: a systematic review and meta-analyses. *Int J Epidemiol* **40**: 65–101.

Hardy L, Jespers V, Van den Bulck M, Buyze J, Mwambarangwe L, Musengamana V, *et al.* (2017). The presence of the putative *Gardnerella vaginalis* sialidase A gene in vaginal specimens is associated with bacterial vaginosis biofilm. *PLoS One* **12**: e0172522.

Harmsen HJ, Wildeboer-Veloo AC, Raangs GC, Wagendorp AA, Klijn N, Bindels JG, *et al.* (2000). Analysis of intestinal flora development in breast-fed and formula-fed infants by using molecular identification and detection methods. *J Pediatr Gastroenterol Nutr* **30**: 61–67.

Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. (2009). Research electronic data

capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* **42**: 377–381.

Hart AL, Lammers K, Brigidi P, Vitali B, Rizzello F, Gionchetti P, *et al.* (2004). Modulation of human dendritic cell phenotype and function by probiotic bacteria. *Gut* **53**: 1602–1609.

Hay PE. (2009). Recurrent bacterial vaginosis. *Curr Opin Infect Dis* **22**: 82–86.

Hay PE, Lamont RF, Taylor-Robinson D, Morgan DJ, Ison C. (1994). Abnormal bacterial colonisation of the genital tract and subsequent preterm delivery and late miscarriage. *Br Med J* **308**: 295–298.

Hernandez-Guerrero C, Monzon-Bordonaba F, Jimenez-Zamudio L, Ahued-Ahued R, Arechavaleta-Velasco F, Strauss III JF, *et al.* (2003). In-vitro secretion of proinflammatory cytokines by human amniochorion carrying hyper-responsive gene polymorphisms of tumour necrosis factor-a and interleukin-1b. *Mol Hum Reprod* **9**: 625–629.

Heumann CL, Quilter LAS, Eastment MC, Heffron R, Hawes SE. (2017). Adverse birth outcomes and maternal *Neisseria gonorrhoeae* infection. *Sex Transm Dis* **44**: 266–271.

Heuvelin E, Lebreton C, Bichara M, Cerf-Bensussan N, Heyman M. (2010). A *Bifidobacterium* probiotic strain and its soluble factors alleviate chloride secretion by human intestinal epithelial cells. *J Nutr* **140**: 7–11.

Hickey R, Abdo Z, Zhou X, Nemeth K, Hansmann M, Osborn T, *et al.* (2013). Effects of tampons and menses on the composition and diversity of vaginal microbial communities over time. *BJOG An Int J Obstet Gynaecol* **120**: 695–706.

Hill DR, Brunner ME, Schmitz DC, Davis CC, Flood JA, Schlievert PM, *et al.* (2005a). In vivo assessment of human vaginal oxygen and carbon dioxide levels during and post menses. *J Appl Physiol* **99**: 1582–1591.

Hill GB. (1993). The microbiology of bacterial vaginosis. *Am J Obstet Gynecol* **169**: 450–454.

Hill JE, Fernando WMU, Zello GA, Tyler RT, Dahl WJ, Van Kessel AG. (2010). Improvement of the representation of bifidobacteria in fecal microbiota metagenomic libraries by application of the cpn60 universal primer cocktail. *Appl Environ Microbiol* **76**: 4550–2.

Hill JE, Goh SH, Money DM, Doyle M, Li A, Crosby WL, *et al.* (2005b). Characterization of vaginal microflora of healthy, nonpregnant women by chaperonin-60 sequence-based methods. *Am J Obstet Gynecol* **193**: 682–692.

Hill JE, Paccagnella A, Law K, Melito PL, Woodward DL, Price L, *et al.* (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. *J Med Microbiol* **55**: 393–399.

Hill JE, Penny SL, Crowell KG, Goh SH, Hemmingsen SM. (2004). cpnDB: a chaperonin

sequence database. *Genome Res* **14**: 1669–1675.

Hill JE, Town JR, Hemmingsen SM. (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. *Environ Microbiol* **8**: 741–746.

Hillier SL. (1993). Diagnostic microbiology of bacterial vaginosis. *Am J Obstet Gynecol* **169**: 455–459.

Hillier SL, Martius J, Krohn M, Kiviat N, Holmes KK, Eschenbach DA. (1988). A case–control study of chorioamnionic infection and histologic chorioamnionitis in prematurity. *N Engl J Med* **319**: 972–978.

Hillier SL, Nugent RP, Eschenbach DA, Krohn MA, Gibbs RS, Martin DH, *et al.* (1995). Association between bacterial vaginosis and preterm delivery of a low-birth-weight infant. *N Engl J Med* **333**: 1737–1742.

Hillier SL, Witkin SS, Krohn MA, Watts DH, Kiviat NB, Eschenbach DA. (1993). The relationship of amniotic fluid cytokines and preterm delivery, amniotic fluid infection, histologic chorioamnionitis, and chorioamnion infection. *Obstet Gynecol* **81**: 941–948.

Hitti J, Garcia P, Totten P, Paul K, Astete S, Holmes KK. (2010). Correlates of cervical *Mycoplasma genitalium* and risk of preterm birth among peruvian women. *Sex Transm Dis* **37**: 81–85.

Holmes K, Spiegel C, Amsel A, Eschenbach D, Chen K, Totten P. (1981). Nonspecific vaginosis. *Scand J Infect Dis* **26**: 110–114.

Holst E, Goffeng AR, Andersch B. (1994). Bacterial vaginosis and vaginal microorganisms in idiopathic premature labor and association with pregnancy outcome. *J Clin Microbiol* **32**: 176–186.

Hooper L or. V., Gordon JI. (2001). Commensal host-bacterial relationships in the gut. *Science* **292**: 1115–1118.

Horvath P, Barrangou R. (2010). CRISPR/Cas, the immune system of Bacteria and Archaea. *Science* **327**: 167–170.

House M, Kaplan DL, Socrate S. (2009). Relationships between mechanical properties and extracellular matrix constituents of the cervical stroma during pregnancy. *Semin Perinatol* **33**: 300–307.

Huang Y-E, Wang Y, He Y, Ji Y, Wang L-P, Sheng H-F, *et al.* (2015). Homogeneity of the vaginal microbiome at the cervix, posterior fornix, and vaginal canal in pregnant chinese women. *Microb Ecol* **69**: 407–414.

Hummelen R, Fernandes AD, Macklaim JM, Dickson RJ, Changalucha J, Gloor GB, *et al.* (2010). Deep sequencing of the vaginal microbiota of women with HIV. *PLoS One* **5**: e12078.

Hunt KM, Foster JA, Forney LJ, Schütte UME, Beck DL, Abdo Z, *et al.* (2011). Characterization of the diversity and temporal stability of bacterial communities in human milk. *PLoS One* **6**: e21313.

Huys G, Vancanneyt M, D'Haene K, Falsen E, Wauters G, Vandamme P. (2007). *Alloscardovia omnicolens* gen. nov., sp. nov., from human clinical samples. *Int J Syst Evol Microbiol* **57**: 1442–1446.

Hyman RW, Fukushima M, Diamond L, Kumm J, Giudice LC, Davis RW. (2005). Microbes on the human vaginal epithelium. *Proc Natl Acad Sci U S A* **102**: 7952–7957.

Hyman RW, Fukushima M, Jiang H, Fung E, Rand L, Johnson B, *et al.* (2014). Diversity of the vaginal microbiome correlates with preterm birth. *Reprod Sci* **21**: 32–40.

Iams JD, Goldenberg RL, Meis PJ, Mercer BM, Moawad A, Das A, *et al.* (1996). The length of the cervix and the risk of spontaneous premature delivery. *N Engl J Med* **334**: 567–573.

Ingianni A, Petruzzelli S, Morandotti G, Pompei R. (1997). Genotypic differentiation of *Gardnerella vaginalis* by amplified ribosomal DNA restriction analysis (ARDRA). *FEMS Immunol Med Microbiol* **18**: 61–66.

Jacobsson B, Ladfors L, Milsom I. (2004). Advanced maternal age and adverse perinatal outcome. *Obstet Gynecol* **104**: 727–733.

Jacobsson B, Mattsby-Baltzer I, Hagberg H. (2005). Interleukin-6 and interleukin-8 in cervical and amniotic fluid: relationship to microbial invasion of the chorioamniotic membranes. *BJOG An Int J Obstet Gynaecol* **112**: 719–724.

Jacobsson B, Pernevi P, Chidekel L, Platz-Christensen JJ. (2002). Bacterial vaginosis in early pregnancy may predispose for preterm birth and postpartum endometritis. *Acta Obstet Gynecol Scand* **81**: 1006–1010.

Jamieson DJ, Theiler RN, Rasmussen SA. (2006). Emerging infections and pregnancy. *Emerg Infect Dis* **12**: 1638–1643.

Jespers V, Menten J, Smet H, Poradosú S, Abdellati S, Verhelst R, *et al.* (2012). Quantification of bacterial species of the vaginal microbiome in different groups of women, using nucleic acid amplification tests. *BMC Microbiol* **12**: 1–10.

Jett BD, Hatter KL, Huycke MM, Gilmore MS. (1997). Simplified agar plate method for quantifying viable bacteria. *Biotechniques* **23**: 648–650.

Johnson HL, Ghanem KG, Zenilman JM, Erbelding EJ. (2011). Sexually transmitted infections and adverse pregnancy outcomes among women attending inner city public sexually transmitted diseases clinics. *Sex Transm Dis* **38**: 167–171.

Juárez Tomás MS, Ocaña VS, Wiese B, Nader-Macías ME, Tomás MSJ, Ocaña VS, *et al.* (2003). Growth and lactic acid production by vaginal *Lactobacillus acidophilus* CRL 1259, and inhibition

of uropathogenic *Escherichia coli*. *J Med Microbiol* **52**: 1117–1124.

Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. (2014). Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *PLoS One* **9**: e104984.

Kaewnopparat S, Dangmanee N, Kaewnopparat N, Srichana T, Chulasiri M, Settharaksa S. (2013). Invitro probiotic properties of *Lactobacillus fermentum* SK5 isolated from vagina ofa healthy woman. *Anaerobe* **22**: 6–13.

Kandler O. (1983). Carbohydrate metabolism in lactic acid bacteria. *Antonie Van Leeuwenhoek* **49**: 209–224.

Kaoutari A El, Armougom F, Gordon JI, Raoult D, Henrissat B. (2013). The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat Rev Microbiol* **11**: 497–504.

Kataoka S, Yamada T, Chou K, Nishida R, Morikawa M, Minami M, *et al.* (2006). Association between preterm birth and vaginal colonization by mycoplasmas in early pregnancy. *J Clin Microbiol* **44**: 51–55.

Kekki M, Kurki T, Pelkonen J, Kukinen-Raty M, Cacciatore B, Paavonen J. (2001). Vaginal clindamycin in preventing preterm birth and peripartal infections in asymptomatic women with bacterial vaginosis: a randomized, controlled trial. *Obstet Gynecol* **97**: 643–648.

Kenyon C, Colebunders R, Crucitti T. (2013). The global epidemiology of bacterial vaginosis: a systematic review. *Am J Obstet Gynecol* **209**: 505–523.

Kharsany AB, Hoosen AA, van den Ende J. (1993). Antimicrobial susceptibilities of *Gardnerella vaginalis*. *Antimicrob Agents Chemother* **37**: 2733–2735.

Kim YM, Bujold E, Chaiworapongsa T, Gomez R, Yoon BH, Thaler HT, *et al.* (2003). Failure of physiologic transformation of the spiral arteries in patients with preterm labor and intact membranes. *Am J Obstet Gynecol* **189**: 1063–1069.

Kindinger LM, Bennett PR, Lee YS, Marchesi JR, Smith A, Cacciatore S, *et al.* (2017). The interaction between vaginal microbiota, cervical length, and vaginal progesterone treatment for preterm birth risk. *Microbiome* **5**: 1–14.

Kistka ZAF, Palomar L, Lee KA, Boslaugh SE, Wangler MF, Cole FS, *et al.* (2007). Racial disparity in the frequency of recurrence of preterm birth. *Am J Obstet Gynecol* **196**: 1–6.

Klatt NR, Cheu R, Birse K, Zevin AS, Perner M, Noël-Romas L, *et al.* (2017). Vaginal bacteria modify HIV tenofovir microbicide efficacy in African women. *Science* **356**: 938–945.

Klebanoff SJ, Hillier SL, Eschenbaeh DA, Waltersdorph AM. (1991). Control of the microbial flora of the vagina by H_2O_2-generating lactobacilli. *J Infect Dis* **164**: 94–100.

Koskiniemi S, Sun S, Berg OG, Andersson DI. (2012). Selection-driven gene loss in bacteria. *PLoS Genet* **8**: e1002787.

Koumans EH, Markowitz LE, Hogan V. (2002). Indications for therapy and treatment recommendations for bacterial vaginosis in nonpregnant and pregnant women: a synthesis of data. *Clin Infect Dis* **35**: S152–S172.

Koumans EH, Sternberg M, Bruce C, McQuillan G, Kendrick J, Sutton M, *et al.* (2007). The prevalence of bacterial vaginosis in the United States, 2001-2004; associations with symptoms, sexual behaviors, and reproductive health. *Sex Transm Dis* **34**: 864–869.

Krohn MA, Hillier SL, Nugent RP, Cotch MF, Carey JC, Gibbs RS, *et al.* (1995). The genital flora of women with intraamniotic infection. *J Infect Dis* **171**: 1475–1480.

van Kuppeveld FJM, van der Logt JTM, Angulo AF, van Zoest MJ, Quint WG V, Niesters HGM, *et al.* (1992). Genus- and species-specific identification of mycoplasmas by 16S rRNA amplification. *Appl Environ Microbiol* **58**: 2606–2615.

Kurdi AM, Mesleh RA, Al-Hakeem MM, Khashoggi TY, Khalifa HM. (2004). Multiple pregnancy and preterm labor. *Saudi Med J* **25**: 632–637.

Kwak D-W, Hwang H-S, Kwon J-Y, Park Y-W, Kim Y-H. (2014). Co-infection with vaginal *Ureaplasma urealyticum* and *Mycoplasma hominis* increases adverse pregnancy outcomes in patients with preterm labor or preterm premature rupture of membranes. *J Matern Neonatal Med* **27**: 333–337.

Lamont RF, Taylor-Robinson D, Wigglesworth JS, Furr PM, Evans RT, Elder MG. (1987). The role of mycoplasmas, ureaplasmas and chlamydiae in the genital tract of women presenting in spontaneous early preterm labour. *J Med Microbiol* **24**: 253–257.

Lan Y, Rosen G, Hershberg R. (2016). Marker genes that are less conserved in their sequences are useful for predicting genome-wide similarity levels between closely related prokaryotic strains. *Microbiome* **4**: 1–13.

Langmead B, Salzberg SL. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.

Larsson P-G, Forsum U. (2005). Bacterial vaginosis - a disturbed bacterial flora and treatment enigma. *APMIS* **113**: 305–316.

Larsson P-G, Platz-Christensen J-J, Dalaker K, Eriksson K, Fåhraeus L, Irminger K, *et al.* (2000). Treatment with 2% clindamycin vaginal cream prior to first trimester surgical abortion to reduce signs of postoperative infection: a prospective, double-blinded, placebo-controlled, multicenter study. *Acta Obstet Gynecol Scand* **79**: 390–396.

Larsson P-G, Platz-Christensen JJ, Thejls H, Forsum U, Påhlson C. (1992). Incidence of pelvic inflammatory disease after first-trimester legal abortion in women with bacterial vaginosis after treatment with metronidazole: a double-blind, randomized study. *Am J Obstet Gynecol* **166**: 100–103.

Lauder AP, Roche AM, Sherrill-Mix S, Bailey A, Laughlin AL, Bittinger K, *et al.* (2016).

Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota. *Microbiome* **4**: 1–11.

Lee DC, Hassan SS, Romero R, Tarca AL, Bhatti G, Gervasi MT, *et al.* (2011a). Protein profiling underscores immunological functions of uterine cervical mucus plug in human pregnancy. *J Proteomics* **74**: 817–828.

Lee DH, Zo YG, Kim SJ. (1996). Nonradioactive method to study genetic profiles of natural bacterial communities by PCR – single-strand-conformation polymorphism. *Appl Environ Microbiol* **62**: 3112–3120.

Lee J-H, O'Sullivan DJ. (2010). Genomic insights into bifidobacteria. *Microbiol Mol Biol Rev* **74**: 378–416.

Lee J, Romero R, Xu Y, Kim J-S, Topping V, Yoo W, *et al.* (2011b). A signature of maternal anti-fetal rejection in spontaneous preterm birth: chronic chorioamnionitis, anti-human leukocyte antigen antibodies, and C4d. *PLoS One* **6**: e16806.

Lee Y, Allport V, Sykes A, Lindstrom T, Slater D, Bennett P. (2003). The effects of labour and of interleukin 1 beta upon the expression of nuclear factor kappa B related proteins in human amnion. *Mol Hum Reprod* **9**: 213–218.

Leitch ECM, Stewart CS. (2002). *Escherichia coli* O157 and non-O157 isolates are more susceptible to L -lactate than to D -lactate. *Appl Environ Microbiol* **68**: 4676–4678.

Leitich H, Bodner-Adler B, Brunbauer M, Kaider A, Egarter C, Husslein P. (2003). Bacterial vaginosis as a risk factor for preterm delivery: a meta-analysis. *Am J Obstet Gynecol* **189**: 139–147.

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

Lewis AL, Lewis WG. (2012). Host sialoglycans and bacterial sialidases: a mucosal perspective. *Cell Microbiol* **14**: 1174–1182.

Lewis ME, Belland RJ, AbdelRahman YM, Beatty WL, Aiyar AA, Zea AH, *et al.* (2014). Morphologic and molecular evaluation of *Chlamydia trachomatis* growth in human endocervix reveals distinct growth patterns. *Front Cell Infect Microbiol* **4**: 1–12.

Lewis WG, Robinson LS, Gilbert NM, Perry JC, Lewis AL. (2013). Degradation, foraging, and depletion of mucus sialoglycans by the vagina-adapted *Actinobacterium Gardnerella vaginalis*. *J Biol Chem* **288**: 12067–12079.

Lewis WG, Robinson LS, Perry J, Bick JL, Peipert JF, Allsworth JE, *et al.* (2012). Hydrolysis of secreted sialoglycoprotein immunoglobulin A (IgA) in *ex vivo* and biochemical models of bacterial vaginosis. *J Biol Chem* **287**: 2079–2089.

Li K, Bihan M, Yooseph S, Methe BA. (2012). Analyses of the microbial diversity across the

human microbiome. *PLoS One* **7**: e32118.

Libby EK, Pascal KE, Mordechai E, Adelson ME, Trama JP. (2008). *Atopobium vaginae* triggers an innate immune response in an in vitro model of bacterial vaginosis. *Microbes Infect* **10**: 439–446.

Ling Z, Kong J, Liu F, Zhu H, Chen X, Wang Y, *et al.* (2010). Molecular analysis of the diversity of vaginal microbiota associated with bacterial vaginosis. *BMC Genomics* **11**: 1–16.

Links MG, Chaban B, Hemmingsen SM, Muirhead K, Hill JE. (2013). mPUMA: a computational approach to microbiota analysis by de novo assembly of operational taxonomic units based on protein-coding barcode sequences. *Microbiome* **1**: 1–7.

Links MG, Demeke T, Gräfenhan T, Hill JE, Hemmingsen SM, Dumonceaux TJ. (2014). Simultaneous profiling of seed-associated bacteria and fungi reveals antagonistic interactions between microorganisms within a shared epiphytic microbiome on *Triticum* and *Brassica* seeds. *New Phytol* **202**: 542–553.

Links MG, Dumonceaux TJ, Hemmingsen SM, Hill JE. (2012). The chaperonin-60 universal target is a barcode for bacteria that enables de novo assembly of metagenomic sequence data. *PLoS One* **7**: e49755.

Lis R, Rowhani-Rahbar A, Manhart LE. (2015). *Mycoplasma genitalium* infection and female reproductive tract disease: a meta-analysis. *Clin Infect Dis* **61**: 418–426.

Liu B, Roberts CL, Clarke M, Jorm L, Hunt J, Ward J. (2013). Chlamydia and gonorrhoea infections and the risk of adverse obstetric outcomes: a retrospective cohort study. *Sex Transm Infect* **89**: 672–678.

Liu L, Johnson HL, Cousens S, Perin J, Scott S, Lawn JE, *et al.* (2012). Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. *Lancet* **379**: 2151–2161.

Llahi-Camp JM, Rai R, Ison C, Regan L, Taylor-Robinson D. (1996). Association of bacterial vaginosis with a history of second trimester miscarriage. *Hum Reprod* **11**: 1575–1578.

Lopes dos Santos Santiago G, Deschaght P, El Aila N, Kiama TN, Verstraelen H, Jefferson KK, *et al.* (2011). *Gardnerella vaginalis* comprises three distinct genotypes of which only two produce sialidase. *Am J Obstet Gynecol* **204**: 450.e1-450.e7.

Lu GC, Schwebke JR, Duffy LB, Cassell GH, Hauth JC, Andrews WW, *et al.* (2001). Midtrimester vaginal *Mycoplasma genitalium* in women with subsequent spontaneous preterm birth. *Am J Obstet Gynecol* **185**: 163–165.

Luo L, Ibaragi T, Maeda M, Nozawa M, Kasahara T, Sakai M, *et al.* (2000). Interleukin-8 levels and granulocyte counts in cervical mucus during pregnancy. *Am J Reprod Immunol* **43**: 78–84.

Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, *et al.* (2012). SOAPdenovo2: an empirically

improved memory-efficient short-read de novo assembler. *Gigascience* **1**: 1–6.

MacIntyre DA, Chandiramani M, Lee YS, Kindinger L, Smith A, Angelopoulos N, *et al.* (2015). The vaginal microbiome during pregnancy and the postpartum period in a European population. *Sci Rep* **5**: 1–9.

Macklaim JM, Fernandes AD, Di Bella JM, Hammond J-A, Reid G, Gloor GB. (2013). Comparative meta-RNA-seq of the vaginal microbiota and differential expression by *Lactobacillus iners* in health and dysbiosis. *Microbiome* **1**: 1–11.

Macones GA, Parry S, Elkousy M, Clothier B, Ural SH, Strauss JF. (2004). A polymorphism in the promoter region of TNF and bacterial vaginosis: preliminary evidence of gene-environment interaction in the etiology of spontaneous preterm birth. *Am J Obstet Gynecol* **190**: 1504–1508.

Makino H, Kushiro A, Ishikawa E, Kubota H, Gawad A, Sakai T, *et al.* (2013). Mother-to-infant transmission of intestinal bifidobacterial strains has an impact on the early development of vaginally delivered infant's microbiota. *PLoS One* **8**: e78331.

Mann JR, McDermott S, Gill T. (2010). Sexually transmitted infection is associated with increased risk of preterm birth in South Carolina women insured by Medicaid. *J Matern Neonatal Med* **23**: 563–568.

Martín-Platero AM, Valdivia E, Maqueda M, 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. *Anal Biochem* **366**: 102–104.

Martin DH, Marrazzo JM. (2016). The vaginal microbiome: current understanding and future directions. *J Infect Dis* **214**: S36–S41.

Martin HL, Richardson BA, Nyange PM, Lavreys L, Hillier SL, Chohan B, *et al.* (1999). Vaginal lactobacilli, microbial flora, and risk of human immunodeficiency virus type 1 and sexually transmitted disease acquisition. *J Infect Dis* **180**: 1863–1868.

Martinez FAC, Balciunas EM, Converti A, Cotter PD, de Souza Oliveira RP. (2013). Bacteriocin production by *Bifidobacterium* spp. A review. *Biotechnol Adv* **31**: 482–488.

Maslowski KM, Mackay CR. (2011). Diet, gut microbiota and immune responses. *Nat Immunol* **12**: 5–9.

Masson L, Maynard C, Brousseau R, Goh SH, Hemmingsen SM, Hill JE, *et al.* (2006). Identification of pathogenic Helicobacter species by chaperonin-60 differentiation on plastic DNA arrays. *Genomics* **87**: 104–112.

Matamoros S, Gras-Leguen C, Le Vacon F, Potel G, De La Cochetiere MF. (2013). Development of intestinal microbiota in infants and its impact on health. *Trends Microbiol* **21**: 167–173.

Mattarelli P, Bonaparte C, Pot B, Biavati B. (2008). Proposal to reclassify the three biotypes of *Bifidobacterium longum* as three subspecies: *Bifidobacterium longum* subsp. *longum* subsp. nov.,

Bifidobacterium longum subsp. *infantis* comb. nov. and *Bifidobacterium longum* subsp. *suis* comb. nov. *Int J Syst Evol Microbiol* **58**: 767–72.

Mckenney EA, Ashwell M, Lambert JE, Fellner V. (2014). Fecal microbial diversity and putative function in captive western lowland gorillas (*Gorilla gorilla gorilla*), common chimpanzees (*Pan troglodytes*), Hamadryas baboons (*Papio hamadryas*) and binturongs (*Arctictis binturong*). *Integr Zool* **9**: 557–569.

McMillan A, Rulisa S, Sumarah M, Macklaim JM, Renaud J, Bisanz JE, *et al.* (2015). A multi-platform metabolomics approach identifies highly specific biomarkers of bacterial diversity in the vagina of pregnant and non-pregnant women. *Sci Rep* **5**: 1–14.

Meis PJ, Goldenberg RL, Mercer B, Moawad A, Das A, McNellis D, *et al.* (1995). The preterm prediction study: significance of vaginal infections. *Am J Obstet Gynecol* **173**: 1231–1235.

Menard JP, Mazouni C, Salem-Cherif I, Fenollar F, Raoult D, Boubli L, *et al.* (2010). High vaginal concentrations of *Atopobium vaginae* and *Gardnerella vaginalis* in women undergoing preterm labor. *Obstet Gynecol* **115**: 134–140.

Mercer BM, Goldenberg RL, Moawad AH, Meis PJ, Iams JD, Das AF, *et al.* (1999). The preterm prediction study: effect of gestational age and cause of preterm birth on subsequent obstetric outcome. *Am J Obstet Gynecol* **181**: 1216–1221.

Metchnikoff É. (1908). The prolongation of life. Optimistic studies. G. P. Putnam's Sons: New York & London.

Meyer D, Stasse-Wolthuis M. (2009). The bifidogenic effect of inulin and oligofructose and its consequences for gut health. *Eur J Clin Nutr* **63**: 1277–1289.

Milani C, Lugli GA, Duranti S, Turroni F, Bottacini F, Mangifesta M, *et al.* (2014). Genomic encyclopedia of type strains of the genus *Bifidobacterium*. *Appl Environ Microbiol* **80**: 6290–6302.

Milani C, Mancabelli L, Lugli GA, Duranti S, Turroni F, Ferrario C, *et al.* (2015). Exploring vertical transmission of bifidobacteria from mother to child. *Appl Environ Microbiol* **81**: 7078–7087.

Mirmonsef P, Hotton AL, Gilbert D, Burgad D, Landay AL, Weber KM, *et al.* (2014). Free glycogen in vaginal fluids is associated with *Lactobacillus* colonization and low vaginal pH. *PLoS One* **9**: e102467.

Mitchell C, Marrazzo J. (2014). Bacterial vaginosis and the cervicovaginal immune response. *Am J Reprod Immunol* **71**: 555–563.

Mitchell MD, Dudley DJ, Edwin SS, Schiller SL. (1991). Interleukin-6 stimulates prostaglandin production by human amnion and decidual cells. *Eur J Pharmacol* **192**: 189–191.

Mitchell MD, Edwin SS, Lundin-Schiller S, Silver RM, Smotkin D, Trautman MS. (1993). Mechanism of interleukin-1β stimulation of human amnion prostaglandin biosynthesis: mediation

via a novel inducible cyclooxygenase. *Placenta* **14**: 615–625.

Mor G, Cardenas I. (2010). The immune system in pregnancy: a unique complexity. *Am J Reprod Immunol* **63**: 425–433.

Morency A, Bujold E. (2007). The effect of second-trimester antibiotic therapy on the rate of preterm birth. *J Obstet Gynaecol Canada* **29**: 35–44.

Morita H, Nakano A, Onoda H, Toh H, Oshima K, Takami H, *et al.* (2011). *Bifidobacterium kashiwanohense* sp. nov., isolated from healthy infant faeces. *Int J Syst Evol Microbiol* **61**: 2610–2615.

Mueller NT, Bakacs E, Combellick J, Grigoryan Z, Dominguez-Bello MG. (2015). The infant microbiome development: mom matters. *Trends Mol Med* **21**: 109–117.

Mwaniki MK, Atieno M, Lawn JE, Newton CRJC. (2012). Long-term neurodevelopmental outcomes after intrauterine and neonatal insults: a systematic review. *Lancet* **379**: 445–452.

Naidu AS, Bidlack WR, Clemens RA. (1999). Probiotic spectra of lactic acid bacteria (LAB). *Crit Rev Food Sci Nutr* **39**: 13–126.

Ness RB, Hillier SL, Kip KE, Soper DE, Stamm CA, McGregor JA, *et al.* (2004). Bacterial vaginosis and risk of pelvic inflammatory disease. *Obstet Gynecol* **104**: 761–769.

Nugent RP, Krohn MA, Hillier SL. (1991). Reability of diagnosing bacterial vaginosis is improved by standardized method of gram stain interpretation. *J Clin Microbiol* **29**: 297–301.

O'Callaghan A, van Sinderen D. (2016). Bifidobacteria and their role as members of the human gut microbiota. *Front Microbiol* **7**: 1–23.

O'Hanlon DE, Lanier BR, Moench TR, Cone RA. (2010). Cervicovaginal fluid and semen block the microbicidal activity of hydrogen peroxide produced by vaginal lactobacilli. *BMC Infect Dis* **10**: 1–8.

O'Hanlon DE, Moench TR, Cone RA. (2011). In vaginal fluid, bacteria associated with bacterial vaginosis can be suppressed with lactic acid but not hydrogen peroxide. *BMC Infect Dis* **11**: 1–8.

O'Hanlon DE, Moench TR, Cone RA. (2013). Vaginal pH and microbicidal lactic acid when lactobacilli dominate the microbiota. *PLoS One* **8**: e80074.

O'Sullivan DJ, Kullen MJ. (1998). Tracking of probiotic bifidobacteria in the intestine. *Int Dairy J* **8**: 513–525.

Oakeshott P, Hay P, Taylor-Robinson D, Hay S, Dohn B, Kerry S, *et al.* (2004). Prevalence of *Mycoplasma genitalium* in early pregnancy and relationship between its presence and pregnancy outcome. *BJOG An Int J Obstet Gynaecol* **111**: 1464–1467.

Oakley BB, Fiedler TL, Marrazzo JM, Fredricks DN. (2008). Diversity of human vaginal bacterial

communities and associations with clinically defined bacterial vaginosis. *Appl Environ Microbiol* **74**: 4898–4909.

Okonechnikov K, Conesa A, García-Alcalde F. (2015). Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**: 292–294.

Oksanen AJ, Blanchet FG, Kindt R, Legendre P, Minchin PR, Hara RBO, *et al.* (2015). Community Ecology Package 'vegan'. R version 2.3-1. Available: https://cran.r-project.org/web/packages/vegan/.

Okun N, Gronau KA, Hannah ME. (2005). Antibiotics for bacterial vaginosis or *Trichomonas vaginalis* in pregnancy: a systematic review. *Obstet Gynecol* **105**: 857–868.

Oliver KL, Hamelin RC, Hintz WE. (2008). Effects of transgenic hybrid aspen overexpressing polyphenol oxidase on rhizosphere diversity. *Appl Environ Microbiol* **74**: 5340–5348.

Olszak T, An D, Zeissig S, Vera MP, Richter J, Franke A, *et al.* (2012). Microbial exposure during early life has persistent effects on natural killer T cell function. *Science* **336**: 489–493.

Onderdonk AB, Zamarchi GR, Walsh JA, Mellor RD, Munõz A, Kass EH. (1986). Methods for quantitative and qualitative evaluation of vaginal microflora during menstruation. *Appl Environ Microbiol* **51**: 333–339.

Orla-Jensen MLD. (1924). La classification des bactéries lactiques. *Lait* **4**: 468–474.

Osset J, Bartolomé RM, García E, Andreu A. (2001). Assessment of the capacity of *Lactobacillus* to inhibit the growth of uropathogens and block their adhesion to vaginal epithelial cells. *J Infect Dis* **183**: 485–491.

Ozer EA, Allen JP, Hauser AR. (2014). Characterization of the core and accessory genomes of *Pseudomonas aeruginosa* using bioinformatic tools Spine and AGEnt. *BMC Genomics* **15**: 1–17.

Pannaraj PS, Li F, Cerini C, Bender JM, Yang S, Rollie A, *et al.* (2017). Association between breast milk bacterial communities and establishment and development of the infant gut microbiome. *JAMA Pediatr* **171**: 647.

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

Paramel Jayaprakash T, Wagner EC, van Schalkwyk J, Albert AYK, Hill JE, Money DM. (2016). High diversity and variability in the vaginal microbiome in women following preterm premature rupture of membranes (PPROM): a prospective cohort study. *PLoS One* **11**: e0166794.

Parry S, Strauss JF. (1998). Premature rupture of the fetal membranes. *N Engl J Med* **338**: 663–670.

Patterson JL, Stull-Lane A, Girerd PH, Jefferson KK. (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**: 392–399.

Peipert JF, Montagno AB, Cooper AS, Sung CJ. (1997). Bacterial vaginosis as a risk factor for upper genital tract infection. *Am J Obstet Gynecol* **177**: 1184–1187.

Penders J, Thijs C, Vink C, Stelma FF, Snijders B, Kummeling I, *et al.* (2006). Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* **118**: 511–521.

Perez-Muñoz ME, Arrieta M-C, Ramer-Tait AE, Walter J. (2017). A critical assessment of the 'sterile womb' and 'in utero colonization' hypotheses: implications for research on the pioneer infant microbiome. *Microbiome* **5**: 1–19.

Perni SC, Vardhana S, Korneeva I, Tuttle SL, Paraskevas LR, Chasen ST, *et al.* (2004). *Mycoplasma hominis* and *Ureaplasma urealyticum* in midtrimester amniotic fluid: association with amniotic fluid cytokine levels and pregnancy outcome. *Am J Obstet Gynecol* **191**: 1382–1386.

Peterson SW, Knox NC, Golding GR, Tyler SD, Tyler AD, Mabon P, *et al.* (2016). A study of the infant nasal microbiome development over the first year of life and in relation to their primary adult caregivers using cpn60 universal target (UT) as a phylogenetic marker. *PLoS One* **11**: e0152493.

Petkau A, Stuart-Edwards M, Stothard P, van Domselaar G. (2010). Interactive microbial genome visualization with GView. *Bioinformatics* **26**: 3125–3126. Available at: https://server.gview.ca.

Petrova MI, Lievens E, Malik S, Imholz N, Lebeer S. (2015). *Lactobacillus* species as biomarkers and agents that can promote various aspects of vaginal health. *Front Physiol* **6**: 1–18.

PHAC - Public Health Agency of Canada. (2014). Report on sexually transmitted infections in Canada: 2013-2014.

Phillippe M, Elovitz M, Saunders T. (2001). Thrombin-stimulated uterine contractions in the pregnant and nonpregnant rat. *J Soc Gynecol Investig* **8**: 260–265.

Phukan N, Parsamand T, Brooks AES, Nguyen TNM, Simoes-Barbosa A. (2013). The adherence of *Trichomonas vaginalis* to host ectocervical cells is influenced by lactobacilli. *Sex Transm Infect* **89**: 455–459.

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

Piot P, Van Dyck E, Totten PA, Holmes KK. (1982). Identification of *Gardnerella (Haemophilus) vaginalis*. *J Clin Microbiol* **15**: 19–24.

Pleckaityte M, Janulaitiene M, Lasickiene R, Zvirbliene A. (2012). Genetic and biochemical diversity of *Gardnerella vaginalis* strains isolated from women with bacterial vaginosis. *FEMS Immunol Med Microbiol* **65**: 69–77.

Poupard JA, Husain I, Norris RF. (1973). Biology of the bifidobacteria. *Bacteriol Rev* **37**: 136–165.

Prince AL, Antony KM, Chu DM, Aagaard KM. (2014). The microbiome, parturition, and timing of birth: more questions than answers. *J Reprod Immunol* **104–105**: 12–19.

Prince AL, Ma J, Kannan PS, Alvarez M, Gisslen T, Harris RA, *et al.* (2016). The placental membrane microbiome is altered among subjects with spontaneous preterm birth with and without chorioamnionitis. *Am J Obstet Gynecol* **214**: 627.e1-627.e16.

Pybus V, Onderdonk AB. (1998). A commensal symbiosis between *Prevotella bivia* and *Peptostreptococcus anaerobius* involves amino acids: potential significance to the pathogenesis of bacterial vaginosis. *FEMS Immunol Med Microbiol* **22**: 317–327.

Quinn J, Munoz FM, Gonik B, Frau L, Cutland C, Mallett-Moore T, *et al.* (2016). Preterm birth: case definition & guidelines for data collection, analysis, and presentation of immunisation safety data. *Vaccine* **34**: 6047–6056.

Rabe LK, Hillier SL. (2003). Optimization of media for detection of hydrogen peroxide production by *Lactobacillus* species. *J Clin Microbiol* **41**: 3260–3264.

Ralph ASG, Rutherford AJ, Wilson JD. (2015). Influence of bacterial vaginosis on conception and miscarriage in the first trimester: cohort study. *Br Med J* **319**: 220–223.

Rautava S, Luoto R, Salminen S, Isolauri E. (2012). Microbial contact during pregnancy, intestinal colonization and human disease. *Nat Rev Gastroenterol Hepatol* **9**: 565–576.

Ravel J, Brotman RM, Gajer P, Ma B, Nandy M, Fadrosh DW, *et al.* (2013). Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. *Microbiome* **1**: 1–6.

Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, *et al.* (2011). Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A* **108**: 4680–4687.

Renthal NE, Williams KC, Mendelson CR. (2013). MicroRNAs—mediators of myometrial contractility during pregnancy and labour. *Nat Rev Endocrinol* **9**: 391–401.

Reuter G. (1971). Designation of type strains for *Bifidobacterium* species. *Int J Syst Bacteriol* **21**: 273–275.

Richter M, Rossello-Mora R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* **106**: 19126–19131.

Richter M, Rossello-Mora R, Oliver Glockner F, Peplies J. (2016). JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* **32**: 929–931.

Roberge S, Villa P, Nicolaides KH, Giguère Y, Vainio M, Bakthi A, *et al.* (2012). Early

administration of low-dose aspirin for the prevention of preterm and term preeclampsia: a systematic review and meta-analysis. *Fetal Diagn Ther* **31**: 141–146.

Rogosa M. (1974). Genus *Bifidobacterium* Orla-Jensen. In: Buchanan R, Gibbons N (eds). *Bergey's Manual of Determinative Bacteriology*. Williams & Wilkins Co.: Baltimore, pp 669–676.

Romero R, Dey SK, Fisher SJ. (2014a). Preterm labor: One syndrome, many causes. *Science* **345**: 760–765.

Romero R, Espinoza J, Kusanovic JP, Gotsch F, Hassan S, Erez O, *et al.* (2006). The preterm parturition syndrome. *BJOG An Int J Obstet Gynaecol* **113**: 17–42.

Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Bieda J, *et al.* (2014b). The vaginal microbiota of pregnant women who subsequently have spontaneous preterm labor and delivery and those with a normal delivery at term. *Microbiome* **2**: 1–15.

Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Nikita L, *et al.* (2014c). The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. *Microbiome* **2**: 1–19.

Romero R, Sirtori M, Oyarzun E, Avila C, Mazor M, Callahan R, *et al.* (1989). Infection and labor V. Prevalence, microbiology, and clinical significance of intraamniotic infection in women with preterm labor and intact membranes. *Am J Obstet Gynecol* **161**: 817–824.

Rours GIJG, De Krijger RR, Ott A, Willemse HFM, De Groot R, Zimmermann LJI, *et al.* (2011). *Chlamydia trachomatis* and placental inflammation in early preterm delivery. *Eur J Epidemiol* **26**: 421–428.

Rozen S, Skaletsky H. (2000). Primer3 on the WWW for general users and for biologist programmers. In: Vol. 132. *Bioinformatics Methods and Protocols*. Humana Press: New Jersey, pp 365–386.

Ruíz FO, Gerbaldo G, García MJ, Giordano W, Pascual L, Barberis IL. (2012). Synergistic effect between two bacteriocin-like inhibitory substances produced by lactobacilli strains with inhibitory activity for *Streptococcus agalactiae*. *Curr Microbiol* **64**: 349–356.

Ryckman KK, Simhan HN, Krohn MA, Williams SM. (2009). Predicting risk of bacterial vaginosis: the role of race, smoking and corticotropin-releasing hormone-related genes. *MHR Basic Sci Reprod Med* **15**: 131–137.

Sabia C, Anacarso I, Bergonzini A, Gargiulo R, Sarti M, Condò C, *et al.* (2014). Detection and partial characterization of a bacteriocin-like substance produced by *Lactobacillus fermentum* CS57 isolated from human vaginal secretions. *Anaerobe* **26**: 41–45.

Sadhu K, Domingue PAG, Chow AW, Nelligan J, Cheng N, Costerton JW. (1989). *Gardnerella vaginalis* has a gram-positive cell-wall ultrastructure and lacks classical cell-wall Iipopolysaccharide. *J Med Microbiol* **29**: 229–235.

Sakata, Kitahara, Sakamoto, Hayashi, Fukuyama, Benno Y. (2002). Unification of *Bifidobacterium infantis* and *Bifidobacterium suis* as *Bifidobacterium longum*. *Int J Syst Evol Microbiol* 1945–1951.

Sakwinska O, Foata F, Berger B, Brüssow H, Combremont S, Mercenier A, *et al.* (2017). Does the maternal vaginal microbiota play a role in seeding the microbiota of neonatal gut and nose? *Benef Microbes* **8**: 763–778.

Satpute SK, Kulkarni GR, Banpurkar AG, Banat IM, Mone NS, Patil RH, *et al.* (2016). Biosurfactant/s from lactobacilli species: properties, challenges and potential biomedical applications. *J Basic Microbiol* **56**: 1140–1158.

Scardovi V, Crociani F. (1974). *Bifidobacterium catenulatum, Bifidobacterium dentium*, and *Bifidobacterium angulatum*: three new species and their deoxyribonucleic acid homology relationships. *Int J Syst Bacteriol* **24**: 6–20.

Scardovi V, Trovatelli L. (1965). The fructose-6-phosphate shunt as peculiar pattern of hexose degradation in the genus *Bifidobacterium*. *Ann Microbiol Enzimol* **15**: 19–29.

Scardovi V, Trovatelli LD, Biavati B, Zani G. (1979). *Bifidobacterium cuniculi, Bifidobacterium choerinum, Bifidobacterium boum*, and *Bifidobacterium pseudocatenulatum*: four new species and their deoxyribonucleic acid homology relationships. *Int J Syst Bacteriol* **29**: 291–311.

Scardovi V, Trovatelli LD, Zani G, Crociani F, Matteuzzi D. (1971). Deoxyribonucleic acid homology relationships among species of the genus *Bifidobacterium*. *Int J Syst Bacteriol* **21**: 276–294.

Schell M, Karmirantzou M, Snel B, Vilanova D, Berger B, Pessi G, *et al.* (2002). The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proc Natl Acad Sci U S A* **99**: 14422–14427.

Schellenberg JJ, Blake Ball T, Lane M, Cheang M, Plummer F. (2008). Flow cytometric quantification of bacteria in vaginal swab samples self-collected by adolescents attending a gynecology clinic. *J Microbiol Methods* **73**: 216–226.

Schellenberg JJ, Dumonceaux TJ, Hill JE, Kimani J, Jaoko W, Wachihi C, *et al.* (2012). Selection, phenotyping and identification of acid and hydrogen peroxide producing bacteria from vaginal samples of Canadian and East African women. *PLoS One* **7**: e41217.

Schellenberg JJ, Jayaprakash TP, Gamage NW, Patterson MH, Vaneechoutte M, Hill JE. (2016). *Gardnerella vaginalis* subgroups defined by cpn60 sequencing and sialidase activity in isolates from Canada, Belgium and Kenya. *PLoS One* **11**: e0146510.

Schellenberg JJ, Links MG, Hill JE, Dumonceaux TJ, Kimani J, Jaoko W, *et al.* (2011). Molecular definition of vaginal microbiota in East African commercial sex workers. *Appl Environ Microbiol* **77**: 4066–4074.

Schellenberg JJ, Links MG, Hill JE, Dumonceaux TJ, Peters GA, Tyler SD, *et al.* (2009).

Pyrosequencing of the chaperonin-60 universal target as a tool for determining microbial community composition. *Appl Environ Microbiol* **75**: 2889–2898.

Schellenberg JJ, Oh AY, Hill JE. (2017). Microbial profiling of cpn 60 universal target sequences in artificial mixtures of vaginal bacteria sampled by nylon swabs or self-sampling devices under different storage conditions. *J Microbiol Methods* **136**: 57–64.

Schempf AH, Branum AM, Lukacs SL, Schoendorf KC. (2007). Maternal age and parity-associated risks of preterm birth: differences by race/ethnicity. *Paediatr Perinat Epidemiol* **21**: 34–43.

Schoeman J, Grové D, Jager M De. (2011). The association between *Chlamydia trachomatis* genital infection and spontaneous preterm labour. *SAJOG South African Med reseach Counc* **12**: 146–149.

Sela DA, Chapman J, Adeuya A, Kim JH, Chen F, Whitehead TR, *et al.* (2008). The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. *Proc Natl Acad Sci U S A* **105**: 18964–18969.

Sender R, Fuchs S, Milo R. (2016). Are we really vastly outnumbered? revisiting the ratio of bacterial to host cells in humans. *Cell* **164**: 337–340.

Shipitsyna E, Roos A, Datcu R, Hallén A, Fredlund H, Jensen JS, *et al.* (2013). Composition of the vaginal microbiota in women of reproductive age - sensitive and specific molecular diagnosis of bacterial vaginosis is possible? *PLoS One* **8**: e60670.

Shynlova O, Lee Y-H, Srikhajon K, Lye SJ. (2013). Physiologic uterine inflammation and labor onset. *Reprod Sci* **20**: 154–167.

Sobel JD. (2009). Antibiotic consideration in bacterial vaginosis. *Curr Infect Dis Rep* **11**: 471–475.

Sobel JD. (2000). Bacterial vaginosis. *Annu Rev Med* **51**: 349–356.

Sobel JD. (1997). Vaginitis. *N Engl J Med* **337**: 1896–1903.

Spiegel CA, Amsel R, Eschenbach D, Schoenknecht F, Holmes KK. (1980). Anaerobic bacteria in nonspecific vaginitis. *N Engl J Med* **303**: 601–607.

Spurbeck RR, Arvidson CG. (2008). Inhibition of *Neisseria gonorrhoeae* epithelial cell interactions by vaginal *Lactobacillus* species. *Infect Immun* **76**: 3124–3130.

Srinivasan S, Morgan MT, Fiedler TL, Djukovic D, Hoffman NG, Raftery D, *et al.* (2015). Metabolic signatures of bacterial vaginosis. *MBio* **6**: 1–16.

Stackebrandt E, Rainey FA, Ward-Rainey NL. (1997). Proposal for a new hierarchic classification system, *Actinobacteria* classis nov. *Int J Syst Bacteriol* **47**: 479–491.

Statistics Canada. (2016). Preterm live births in Canada, 2000 to 2013. Health Fact Sheets. https://www.statcan.gc.ca/pub/82-625-x/2016001/article/14675-eng.htm.

Steel JH, Malatos S, Kennea N, Edwards AD, Miles L, Duggan P, *et al.* (2005). Bacteria and inflammatory cells in fetal membranes do not always cause preterm labor. *Pediatr Res* **57**: 404–411.

Stock S, Norman J. (2010). Preterm and term labour in multiple pregnancies. *Semin Fetal Neonatal Med* **15**: 336–341.

Stout MJ, Zhou Y, Wylie KM, Tarr PI, Macones GA, Tuuli MG. (2017). Early pregnancy vaginal microbiome trends and preterm birth. *Am J Obstet Gynecol* **217**: 356.e1-356.e18.

Stoyancheva G, Marzotto M, Dellaglio F, Torriani S. (2014). Bacteriocin production and gene sequencing analysis from vaginal *Lactobacillus* strains. *Arch Microbiol* **196**: 645–653.

Strus M, Brzychczy-Włoch M, Gosiewski T, Kochan P, Heczko PB. (2006). The in vitro effect of hydrogen peroxide on vaginal microbial communities. *FEMS Immunol Med Microbiol* **48**: 56–63.

Sugahara H, Odamaki T, Fukuda S, Kato T, Xiao J, Abe F, *et al.* (2015). Probiotic *Bifidobacterium longum* alters gut luminal metabolism through modification of the gut microbial community. *Nature* **5**: 1–11.

Sun Z, Zhang W, Guo C, Yang X, Liu W, Wu Y, *et al.* (2015). Comparative genomic analysis of 45 type strains of the genus *Bifidobacterium*: a snapshot of its genetic diversity and evolution. *PLoS One* **10**: e0117912.

Sundquist A, Bigdeli S, Jalili R, Druzin ML, Waller S, Pullen KM, *et al.* (2007). Bacterial flora-typing with targeted, chip-based pyrosequencing. *BMC Microbiol* **7**: 1–11.

Swidsinski A, Mendling W, Loening-Baucke V, Ladhoff A, Swidsinski S, Hale LP, *et al.* (2005). Adherent biofilms in bacterial vaginosis. *Obstet Gynecol* **106**: 1013–1023.

Swidsinski A, Mendling W, Loening-Baucke V, Swidsinski S, Dörffel Y, Scholze J, *et al.* (2008). An adherent *Gardnerella vaginalis* biofilm persists on the vaginal epithelium after standard therapy with oral metronidazole. *Am J Obstet Gynecol* **198**: 1–6.

Taha TE, Hoover DR, Dallabetta GA, Kumwenda NI, Mtimavalye LAR, Yang L-P, *et al.* (1998). Bacterial vaginosis and disturbances of vaginal flora. *AIDS* **12**: 1699–1706.

Takahashi H, Mikami K, Nishino R, Matsuoka T, Kimura M, Koga Y. (2010). Comparative analysis of the properties of bifidobacterial isolates from fecal samples of mother–infant pairs. *J Pediatr Gastroenterol Nutr* **51**: 653–660.

Taylor-Robinson D. (2017). Mollicutes in vaginal microbiology: *Mycoplasma hominis, Ureaplasma urealyticum, Ureaplasma parvum* and *Mycoplasma genitalium*. *Res Microbiol* 1–7.

Taylor-Robinson D. (2007). The role of mycoplasmas in pregnancy outcome. *Best Pract Res Clin*

Obstet Gynaecol **21**: 425–438.

Taylor-Robinson D, Jensen JS. (2010). Genital Mycoplasmas. In: *Atlas of Sexually Transmitted Diseases and AIDS*. Elsevier, pp 64–75.

Taylor-Robinson D, Lamont R. (2011). Mycoplasmas in pregnancy. *BJOG An Int J Obstet Gynaecol* **118**: 164–174.

Thies FL, Konig W, Konig B. (2007). Rapid characterization of the normal and disturbed vaginal microbiota by application of 16S rRNA gene terminal RFLP fingerprinting. *J Med Microbiol* **56**: 755–761.

Thinkhamrop J, Hofmeyr GJ, Adetoro O, Lumbiganon P, Ota E. (2015). Antibiotic prophylaxis during the second and third trimester to reduce adverse pregnancy outcomes and morbidity. In: Thinkhamrop J (ed). *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd: Chichester, UK, pp 1–35.

Thompson JD, Higgins DG, Gibson TJ. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.

Thomson AJ, Telfer JF, Young A, Campbell S, Stewart CJR, Cameron IT, *et al.* (1999). Leukocytes infiltrate the myometrium during human parturition: further evidence that labour is an inflammatory process. *Hum Reprod* **14**: 229–236.

Tissier H. (1889). La reaction chromophile d'Escherichel et le *Bacterium coli*. *Crit Rev Acad Sci* **51**: 943—945.

Tissier H. (1900). Recherches sur la flore intestinale des nourrissons (etat normal et pathologique). University of Paris.

Tita ATN, Andrews WW. (2010). Diagnosis and management of clinical chorioamnionitis. *Clin Perinatol* **37**: 339–354.

Trama JP, Pascal KE, Zimmerman J, Self MJ, Mordechai E, Adelson ME. (2008). Rapid detection of *Atopobium vaginae* and association with organisms implicated in bacterial vaginosis. *Mol Cell Probes* **22**: 96–102.

Tulchinsky D, Hobel CJ, Yeager E, Marshall JR. (1972). Plasma estrone, estradiol, estriol, progesterone, and 17-hydroxyprogesterone in human pregnancy. *Am J Obstet Gynecol* **112**: 1095–1100.

Turroni F, Bottacini F, Foroni E, Mulder I, Kim J-H, Zomer A, *et al.* (2010). Genome analysis of *Bifidobacterium bifidum* PRL2010 reveals metabolic pathways for host-derived glycan foraging. *Proc Natl Acad Sci U S A* **107**: 19514–19519.

Turroni F, Peano C, Pass DA, Foroni E, Severgnini M, Claesson MJ, *et al.* (2012). Diversity of bifidobacteria within the infant gut microbiota. *PLoS One* **7**: e36957.

Turroni F, Ribbera A, Foroni E, van Sinderen D, Ventura M. (2008). Human gut microbiota and bifidobacteria: from composition to functionality. *Antonie Van Leeuwenhoek* **94**: 35–50.

Ventura M, Canchaya C, Casale AD, Dellaglio F, Neviani E, Fitzgerald GF, *et al.* (2006). Analysis of bifidobacterial evolution using a multilocus approach. *Int J Syst Evol Microbiol* **56**: 2783–2792.

Ventura M, Canchaya C, Tauch A, Chandra G, Fitzgerald GF, Chater KF, *et al.* (2007). Genomics of *Actinobacteria*: tracing the evolutionary history of an ancient phylum. *Microbiol Mol Biol Rev* **71**: 495–548.

Ventura M, O'Flaherty S, Claesson MJ, Turroni F, Klaenhammer TR, van Sinderen D, *et al.* (2009a). Genome-scale analyses of health-promoting bacteria: probiogenomics. *Nat Rev Microbiol* **7**: 61–71.

Ventura M, Turroni F, Zomer A, Foroni E, Giubellini V, Bottacini F, *et al.* (2009b). The *Bifidobacterium dentium* Bd1 genome sequence reflects its genetic adaptation to the human oral cavity. *PLoS Genet* **5**: e1000785.

Verbeke TJ, Sparling R, Hill JE, Links MG, Levin D, Dumonceaux TJ. (2011). Predicting relatedness of bacterial genomes using the chaperonin-60 universal target (cpn60 UT): application to *Thermoanaerobacter* species. *Syst Appl Microbiol* **34**: 171–179.

Verhelst R, Verstraelen H, Claeys G, Verschraegen G, Van Simaey L, De Ganck C, *et al.* (2005). Comparison between Gram stain and culture for the characterization of vaginal microflora: definition of a distinct grade that resembles grade I microflora and revised categorization of grade I microflora. *BMC Microbiol* **5**: 1–11.

Vermette CJ, Russell AH, Desai AR, Hill JE. (2010). Resolution of phenotypically distinct strains of *Enterococcus* spp. in a complex microbial community using *cpn*60 universal target sequencing. *Microb Ecol* **59**: 14–24.

Verstraelen H, Verhelst R. (2009). Bacterial vaginosis: an update on diagnosis and treatment. *Expert Rev Anti Infect Ther* **7**: 1109–1124.

Verstraelen H, Verhelst R, Claeys G, De Backer E, Temmerman M, 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 vagin. *BMC Microbiol* **9**: 1–10.

de Vries W, Gerbrandy SJ, Stouthamer AH. (1967). Carbohydrate metabolism in *Bifidobacterium bifidum*. *Biochim Biophys Acta - Gen Subj* **136**: 415–425.

Walencka E, Różalska S, Sadowska B, Różalska B. (2008). The influence of *Lactobacillus acidophilus*-derived surfactants on staphylococcal adhesion and biofilm formation. *Folia Microbiol (Praha)* **53**: 61–66.

Walker AW. (2016). Studying the Human Microbiota. In: Schwiertz A (ed) Advances in Experimental Medicine and Biology Vol. 902. *Microbiota of the Human Body*. Springer: Cham,

pp 5–32.

Walker AW, Martin JC, Scott P, Parkhill J, Flint HJ, Scott KP. (2015). 16S rRNA gene-based profiling of the human infant gut microbiota is strongly influenced by sample processing and PCR primer choice. *Microbiome* **3**: 1–11.

Walther-António MRS, Jeraldo P, Berg Miller ME, Yeoman CJ, Nelson KE, Wilson BA, *et al.* (2014). Pregnancy's stronghold on the vaginal microbiome. *PLoS One* **9**: e98514.

Watson HL, Blalock DK, Cassell GH. (1990). Variable antigens of *Ureaplasma urealyticum* containing both serovar-specific and serovar-cross-reactive epitopes. *Infect Immun* **58**: 3679–3688.

Watts DH, Krohn MA, Hillier SL, Eschenbach DA. (1990). Bacterial vaginosis as a risk factor for post-cesarean endometritis. *Obstet Gynecol* **75**: 52–58.

Watts HD, Krohn MA, Hillier SL, Eschenbach DA. (1992). The association of occult amniotic fluid infection with gestational age and neonatal outcome among women in preterm labor. *Obstet Gynecol* **79**: 351–357.

WHO, March of Dimes, PMNCH, Save the Children. (2012). Born too soon: the global action report on preterm birth. In: Howson CP, Kinney M V, Lawn JE (eds). World Health Organization: Geneva.

Wi T, Lahra MM, Ndowa F, Bala M, Dillon JR, Ramon-Pardo P, *et al.* (2017). Antimicrobial resistance in *Neisseria gonorrhoeae*: global surveillance and a call for international collaborative action. *PLoS Med* **14**: e1002344.

Wiesenfeld HC, Hillier SL, Krohn MA, Landers D V, Sweet RL. (2003). Bacterial vaginosis is a strong predictor of *Neisseria gonorrhoeae* and *Chlamydia trachomatis* infection. *Clin Infect Dis* **36**: 663–668.

Wilks M, Tabaqchali S. (1987). Quantitative bacteriology of the vaginal flora during the menstrual cycle. *J Med Microbiol* **24**: 241–245.

Witkin SS, Mendes-Soares H, Linhares IM, Jayaram A, Ledger WJ, Forney LJ. (2013). Influence of vaginal bacteria and D- and L-lactic acid isomers on vaginal extracellular matrix metalloproteinase inducer: Implications for protection against upper genital tract infections. *MBio* **4**: 1–7.

Xiong X, Buekens P, Fraser W, Beck J, Offenbacher S. (2006). Periodontal disease and adverse pregnancy outcomes: a systematic review. *BJOG An Int J Obstet Gynaecol* **113**: 135–143.

Yang SW, Li W, Challis JRG, Reid G, Gloor GB, Seney SL, *et al.* (2015). Oral *Lactobacillus rhamnosus* GR-1/*L. reuteri* RV-14 does not change the vaginal microbiome or cervico-vaginal cytokines in low risk pregnant women with an abnormal Nugent score. *Reprod Sci* **22**: 90A.

Yatsunenko T, Rey FE, Manary MJMMJ, Trehan I, Dominguez-Bello MG, Contreras M, *et al.*

(2012). Human gut microbiome viewed across age and geography. *Nature* **486**: 222–227.

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13**: 1–11.

Yoon BH, Romero R, Moon J Bin, Shim S-S, Kim M, Kim G, *et al.* (2001). Clinical significance of intra-amniotic inflammation in patients with preterm labor and intact membranes. *Am J Obstet Gynecol* **185**: 1130–1136.

Yoon BH, Romero R, Park JS, Chang JW, Kim YA, Kim JC, *et al.* (1998). Microbial invasion of the amniotic cavity with *Ureaplasma urealyticum* is associated with a robust host response in fetal, amniotic, and maternal compartments. *Am J Obstet Gynecol* **179**: 1254–1260.

Zevin AS, Xie IY, Birse K, Arnold K, Romas L, Westmacott G, *et al.* (2016). Microbiome composition and function drives wound-healing impairment in the female genital tract. *PLoS Pathog* **12**: e1005889.

Zhou X, Brotman RM, Gajer P, Abdo Z, Schütte UME, Ma S, *et al.* (2010). Recent advances in understanding the microbiology of the female reproductive tract and the causes of premature birth. *Infect Dis Obstet Gynecol* 1–10.

Zhou X, Brown CJ, Abdo Z, Davis CC, Hansmann MA, Joyce P, *et al.* (2007). Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women. *ISME J* **112**: 121–133.

Ziklo N, Huston WM, Hocking JS, Timms P. (2016). *Chlamydia trachomatis* genital tract infections: when host immune response and the microbiome collide. *Trends Microbiol* **24**: 750–765.

Zinedine A, Faid M. (2007). Isolation and characterization of strains of bifidobacteria with probiotic proprieties in vitro. *World J Dairy Food Sci* **2**: 28–34.

Zinnemann K, Turner G. (1963). The taxonomic position of *'Haemophilus vaginalis' [Corynebacterium vaginale]*. *J Pathol Bacteriol* **85**: 213–219.