A METAGENOMIC APPROACH TO UNDERSTANDING THE RELATIONSHIP BETWEEN THE VAGINAL MICROBIOME AND WOMEN'S REPRODUCTIVE HEALTH

BY

FANG YANG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Nutritional Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Doctoral Committee:

Professor Kelly Swanson, Chair
Professor Lawrence B. Schook, Director of Research
Professor Bryan A. White, Co-Director of Research
Associate Professor Saurabh Sinha
Assistant Professor Nicholas Chia

ABSTRACT

Vaginal microbes exert a strong influence on women's reproductive health. Perturbations in the vaginal microbiome have been associated with serious sequelae, such as bacterial vaginosis (BV) and pregnancy complications. The etiology of BV is multifactorial. To gain a mechanistic understanding of BV, there is a need to understand what pathogens initiate it and how they do so. However, a minority of existing research on BV focused on the latter. Sialidase assists the colonization of pathogens by destroying mucin. Therefore, we hypothesized that sialidase could serve as a biomarker for BV (chapter 2). We used the BVBlue sialidase test to measure the sialidase activity, 454-pyrosequencing of amplicons from the V1-V3 region of the 16S rRNA gene to classify sequence to bacteria, and Illumina high-throughput sequencing to classify sequence to bacteria and functional genes in microbial samples from 38 Chinese women with and without BV from 20 to 52 years (means ± SEM 35.6 ± 1.2 years). The samples were collected by swabs from cervix, middle, and upper vagina walls. Elevated sialidase activity (>= 7.8U) was detected in all BV patients but not in any Chinese women without BV. Sequence-based analysis indicated that *Prevotella* and *Gardnerella* species were more abundant in Chinese BV patients by linear discriminant analysis (LDA) effect size (LEfSe). The conclusion was that the sialidase test could separate diseased vaginal microbiomes from non-diseased ones. *Prevotella* and *Gardnerella* species were associated with high sialidase levels.

An alternative way to study how pathogens initiate BV is to study how pathogens are involved in the production of BV symptoms. Yeoman *et al.* [1] found vaginal bacteria and metabolites that are correlated with BV symptoms. Therefore, we hypothesized that BV symptom-associated bacteria could serve as potential biomarkers to improve the development of therapeutic targets (chapter

3). To better understand how pathogens are involved in the production of BV symptoms, we took a subset of the population (n=4) from Yeoman *et al.* [1]. All women were of reproductive age from 20 to 40 years. The samples were collected by swabs, scrapings, and lavage. 454-pyrosequencing of amplicons from the V1-V3 region of the 16S rRNA gene was used to classify sequence to bacteria. Illumina high-throughput sequencing was used to classify sequence to bacteria and functional genes. We then used Spearman's rank correlation coefficients to examine the correlation among vaginal bacteria, vaginal discharge-associated metabolites, and functional genes. The bacteria *Aquabacterium*, *Chryseobacterium*, *Elizabethkingia*, *Finegoldia*, *Morganella*, *Negativicoccus*, *Phenylobacterium*, *Propionibacterium*, *Serratia*, *Siphonobacter*, and the enzymes 'malate dehydrogenase' (K00024), '5-amino-6-(5-phosphoribosylamino) uracil reductase' (K11752), and 'undecaprenyl-diphosphatase' (K06153) were positively associated with 2-Methyl-2-hydroxybutanoic acid (r=1, p=0) that was the only metabolite revealed by Yeoman *et al.* [1] that positively correlated with discharge. Moreover, these three enzymes were positively linked to these ten bacterial genera (r=1, p=0). We conclude that the bacteria *Aquabacterium*, *Chryseobacterium*, *Elizabethkingia*, *Finegoldia*, *Morganella*, *Negativicoccus*, *Phenylobacterium*, *Propionibacterium*, *Serratia*, and *Siphonobacter* are potential therapeutic targets.

Healthy women's vaginas are mainly dominated by *Lactobacilli* at reproductive age. The predominance of *Lactobacilli* is coincident with the rise in estrogen levels and glycogen content in the vaginal epithelium. Estrogen levels increase gradually during pregnancy and drop down at post-partum. Glycogen is converted to lactic acid mainly by *Lactobacilli*. Therefore, we hypothesized that the abundance of *Lactobacilli* and functional genes for metabolic activity in

Lactobacilli would increase as pregnancy progresses and decrease following parturition (chapter 4). Vaginal samples were collected from 5 Caucasian women with and without a history of preterm birth at 8–12, 17–21, 26–30, and 35–38 weeks of gestation, during labor, and at 6 weeks post-partum. All women lived in Rochester, MN, USA from 23 to 43 years. The samples were collected by swabs from the posterior fornix and cervix. We used a metagenomic approach to classify sequences to bacterial species, GroopM [102] and CheckM [104] to recover the *Lactobacillus* population genomes, and DIAMOND [105] to blast the predicted genes in the *Lactobacillus* population genomes against the KEGG database (v56) to classify sequences to metabolic pathways. Generally, the majority of vaginal bacteria (defined as more than half of the vaginal bacteria) were *Lactobacillus* species in women who delivered at term and who did not have a previous preterm birth history during pregnancy. However, the trend was not seen in women who had a previous preterm birth history but still delivered at term. The abundance of *Lactobacilli* and functional genes for metabolic activity in *Lactobacilli* were higher in the samples collected in late pregnancy than in the samples collected at post-partum in all women with term birth. The abundance of *Lactobacilli* and functional genes for metabolic activity in *Lactobacilli* did not increase in any women as pregnancy progressed from 8 weeks of gestational age to 30 weeks of gestational age. We conclude that the vaginal microbiomes of women who delivered at term and who did not have a previous preterm birth history were generally dominated by *Lactobacillus* species during pregnancy. The vaginal microbiomes of women who delivered at term but had a previous preterm birth history were more diverse than those of women who delivered at term and who did not have a previous preterm birth history.

ACKNOWLEDGEMENTS

This dissertation research would not have been possible without the support of many people. First, I would like to thank my two Ph.D. advisors, Drs. Lawrence B. Schook and Bryan A. White for their guidance and support throughout my Ph.D. career. Drs. Lawrence B. Schook and Bryan A. White have taught me how to be a good scientist. Second, I would like to thank my committee members, Drs. Kelly Swanson, Nicholas Chia, and Saurabh Sinha for their wonderful suggestions over the years. Third, I would like to thank people from Dr. Schook's lab, Dr. Laurie Rund, Dr. Kyle Schachtschneider, Dr. Radhika Ganu, Dr. Ying Liang, Dr. Maria Pieters, Dr. Kwame Darfour-Oduro, Dr. Molly Melhem, and Arun Kumar De, people from Dr. White's lab, Dr. Andrea Braundmeier-Fleming, Dr. Margret Miller, Dr. Melissa Cregger, Dr. Carl Yeoman, Dr. Nam-phuong Nguyen and Dr. Andres Gomez, and people from Dr. Chia's lab, Dr. Marina Walther-Antonio, for their help during my career as a Ph.D. student. Fourth, I would like to thank people from the Carl R. Woese Institute for Genomic Biology, Dr. Chris Fields, Dr. Yun-juan Chang, and David Slater for their technical support. I would like to thank Dr. Matthew Benedict for his critical editing contributions. Lastly, I would like to thank my parents, parents-in-law, husband, and son as well as friends for their encouragement and support throughout my graduate career.

TABLE OF CONTENTS

CHAPTER 1. LITERATURE REVIEW

1.1 The human microbiome

1.1.1 Living in a microbial world

We live in a microbial world. Humans are mainly composed of microorganisms. Human-associated microbes outnumber human cells by 10:1 [2,3]. The genes of human-associated microbes outnumber human genes by 100:1 [3]. Microorganisms inhabit almost every part of the human body [4,5]. There are microbes in the nasal and oral cavities, on the skin, in the gut and in the urogenital tract [4,5]. The oral cavities, skin, gut, and urogenital tract each harbor distinct bacteria [5]. Bacterial diversity was highest in oral cavities and lowest in urogenital tract [5]. Although no one bacteria is present in all samples of any one body site, most people's oral cavities, skin, gut, and vagina are dominated by *Streptococcus*, *Propionibacterium*, *Bacteroides*, and *Lactobacillus*, respectively [5].

1.1.2 Symbiotic relationship between microorganisms and humans

Human-associated microbes not only live in and on human bodies, but also form symbiotic relationships with humans. There are three types of symbiosis: mutualism, commensalism, and parasitism. Mutualism is the win-win type of symbiosis where two organisms benefit from collaboration. For instance, the host provides nutrients, like glycogen, to vaginal microorganisms to support their growth. In turn, vaginal microorganisms inhibit the colonization of pathogens by producing antimicrobial compounds, such as lactic acid, hydrogen peroxide and bacteriocins. The host provides nutrients, like dietary fiber, to gut microorganisms to support their growth. In turn, gut

microorganisms provide vitamins to human beings. Commensalism is the symbiosis where one organism is benefited while the other is not impacted. Parasitism is the symbiosis where one organism benefits at the cost of the other. All pathogens are parasites. Disease occurs when homeostasis is disturbed.

Human-associated microbes exert a strong influence on human health and diseases. However, the relationship between microorganisms and human health is poorly understood. In this study, I focus on the relationship between the vaginal microbiome and women's reproductive health.

1.2 The vaginal microbiome

1.2.1 The composition of the vaginal microbiome

Lactobacilli have been thought to be the keystone vaginal species for a long time because they can generate lactic acid to maintain vagina acidity as well as hydrogen peroxide and bacteriocins to kill pathogens in the vagina [6,7]. This point of view is rooted in a previous study dating from 1892, in which Döderlein isolated *Lactobacilli* from vaginal discharge using a culture-dependent method and found that the bacteria could produce lactic acid to inhibit the colonization of pathogens both *in vitro* and *in vivo* [6,8].

With the advances in culture-independent methods, researchers found that *Lactobacilli* were not the keystone vaginal species for all healthy women. Researchers used 16S-rRNA based approaches to disclose the profile of healthy vaginal microbiomes and found that *Lactobacillus* species dominated in a majority of healthy women's vaginas [7].

However, appreciable amounts of healthy women had vaginal microbiome predominated by other various anaerobic microorganisms [7]. For example, Ravel *et al.* [9] studied the vaginal microbiomes of 396 non-pregnant healthy American women representing four ethnic groups (white, black, Hispanic, and Asian). They showed that vaginal microbial community composition varies among non-pregnant healthy women [9]. Non-pregnant women's vaginal microbiomes could be clustered into five different community state types (CSTs) [9]. Four out of five CSTs, representing 73% of the non-pregnant healthy women's vaginal microbiomes, are dominated by one species of *Lactobacillus* [9]. The remaining CST is not dominated by *Lactobacillus* or by any one bacterium [9]. Instead, it is dominated by various anaerobes [9]. The same research group found that healthy pregnant women's vaginal microbiome could be clustered into four different CSTs [10]. Three out of four CSTs are dominated by one species of *Lactobacillus* [10]. The remaining CST is not dominated by *Lactobacillus* or by any one bacterium [10].

1.2.2 The effect of estrogen on the vaginal microbiome composition

Researchers also found that *Lactobacilli* were not always the keystone vaginal species for the same woman over a lifetime [6]. The vaginal microbiome composition changes over a woman's lifetime and is related to the level of estrogen [6,11]. Hickey *et al.* [6] and Farage *et al.* [11] showed that in early childhood, the estrogen levels are low. Low estrogen levels lead to the thinning of the vaginal epithelium, a decrease of glycogen content in the vaginal epithelium, a nearly neutral vaginal pH value, and poly-microbial communities in the vagina [6,11]. During puberty and reproductive years, estrogen production rises, resulting in an increase in the thickness of the vaginal epithelium, an increase of glycogen

content in the vaginal epithelium, an acidic vaginal pH value, and the predominance of *Lactobacilli* in the vagina [6,11]. At postmenopause, estrogen levels decrease, causing a decrease of glycogen content in the vaginal epithelium and a nearly neutral vaginal pH value [6,11]. The vaginal microbiome shifts from a *Lactobacilli* dominated niche to poly-microbial communities during this time [6,11]. The vaginal microbiome composition was also reported to change throughout the menstrual cycle [3,12]. At menses, estrogen levels are at their lowest in the menstrual cycle and non-*Lactobacillus* species increase compared with other times in the menstrual cycle [12]. The vaginal microbiome composition was reported to differ during pregnancy compared to when not pregnant [10,13,14]. Vaginal microbiomes during pregnancy were reported to have more *Lactobacillus* species [10,13-15]. The change in the vaginal microbiome composition during pregnancy is probably caused by increases in estrogen levels during pregnancy. For example, the estrogen levels during pregnancy are 100-1000 fold higher than those at 1-week postpartum [13,16,17]. High estrogen levels increase glycogen content in the vaginal epithelium and promote the growth of *Lactobacillus* species [11,18].

So far, only two studies longitudinally examined the change in the vaginal microbiomes during pregnancy [13,14]. Both of them only studied the change in the vaginal microbiome taxonomic composition during pregnancy [13,14]. They did not recover the *Lactobacillus* genome or determine how functional composition of vaginal microbiome changes during pregnancy.

1.2.3 The effect of race and hygiene on the vaginal microbiome composition

The above paragraph shows how estrogen influences the vaginal microbiome composition. The vaginal microbiome composition is also affected by race and hygiene. For instance, Ravel *et al.* [9] revealed that *Lactobacillus* is the predominant vaginal bacterial species in 80% of non-pregnant healthy Asian American women, 90% of non-pregnant healthy Caucasian women, 60% of non-pregnant healthy African American women, and 62% of non-pregnant healthy Hispanic women [9]. Various anaerobes dominated in the vagina of 20% of non-pregnant healthy Asian American women, 10% of non-pregnant healthy Caucasian women, 40% of non-pregnant healthy African American women, and 38 % of non-pregnant healthy Hispanic women [9]. When compared to non-pregnant healthy African American and Hispanic women, non-pregnant healthy Caucasian and Asian American women were more likely to have vaginal communities dominated by *Lactobacillus* [9]. Fettweis *et al.* [19] revealed that there are significant differences in the vaginal microbiome composition between African American women and women of European ancestry. Bacteria associated with bacterial vaginosis were more commonly found in the vaginas of African American women compared with women of European ancestry [19].

Douching has also been to be reported to change the vaginal microbiome composition [20,21]. For example, Onderdonk *et al.* [20] revealed that the use of acetic acid and saline douches significantly reduce the total bacterial counts within ten minutes.

1.3 Women's reproductive health

1.3.1 Bacterial vaginosis

BV is one of the most common, major problems impacting women's reproductive health worldwide [22]. The prevalence of BV in the United States is about 29.2% or approximately 21.2 million women of reproductive age [23] and is associated with serious sequelae. In non-pregnant women, BV has been linked predominantly to the upper genital tract and poses an increased risk for sexually transmitted infections [3,24-27] and susceptibility to human immunodeficiency virus (HIV) [27-29]. In pregnant women, BV has been linked to miscarriages [27,30-32], post-abortion infections [27,33], postpartum endometritis [27,34,35], preterm pre-labor rupture of membranes (PPROM) [36], and preterm birth [3,27,37,38]. The United States annually spends nearly $1 billion on treating BV-related pregnancy complications [39].

The etiology of BV is multifactorial. BV is often ambiguously characterized as perturbations in normal vaginal microbiome [7], and believed is a result of a shift from a *Lactobacilli*-dominated to a poly-microbial community [38]. For example, BV patients are believed to have less *Lactobacilli* and more *Gardnerella* and *Mobiluncus* compared with healthy asymptomatic women. The potential mechanisms supporting the ambiguous characterization of BV are: (1) *Lactobacillus* species produce lactic acid to maintain an acidic vaginal environment (pH < 4.5), hydrogen peroxide and bacteriocin to kill pathogens in vagina; and (2) anaerobic bacteria produce polyamines that cause a "fishy" odor in the vaginal discharge, as well as mucinases and sialidases enzymes, which destroy mucosal tissue and facilitate colonization of pathogens in the vagina [40]. However,

recent studies suggest that *Lactobacilli* do not always predominate in the vaginal microbiomes of healthy women. In 10 to 42% of healthy asymptomatic women, *Lactobacilli* were outnumbered by other vaginal microbes, such as *Atopobium vaginae*, *Leptotrichia*, and *Megasphaera* [9,41-48]. Species like *Gardnerella vaginalis* and *Mobiluncus curtisii*, which have been traditionally associated with BV, are also commonly found in healthy asymptomatic women [49]. Although the exact cause of BV is still unknown, *Gardnerella vaginalis*, *Atopobium*, *Prevotella*, *Porphyromonas*, *Peptostreptococcus*, *Mobiluncus*, *Mycoplasma*, *Eggerthella*-like bacterium, *Megasphaera*, *Leptotrichia*, *Dialister*, and several BV associated bacteria (BVABs) from the Clostridiales order were generally believed to be associated with BV [50,51].

Gradually, researchers found that a simple understanding of the taxonomic composition of the vaginal microbiomes of BV patients could not comprehensively describe the exact mechanism of BV. For example, it has been reported that the genome of *Gardnerella vaginalis* strains that were isolated from BV patients encode for mucin degradation enzymes [52]. However, the genomes of *Gardnerella vaginalis* strains isolated from healthy women did not encode these genes [52]. The degradation of mucin facilitates colonization of pathogens and initiation of BV. Therefore, to gain a mechanistic understanding of BV, there is a need to understand what pathogens initiate it and how they do so. However, a minority of existing research on BV focused on answering the question of how pathogens initiate BV. Both Yeoman *et al.* [1] and Srinivasan *et al.* [53] revealed that women with and without BV have distinct metabolite profiles and found metabolites that are correlated with BV symptoms: discharge and odor. Yeoman *et al.* [1] found that BV discharge is

associated with 2-methyl-2-hydroxybutanoic acid that were linked to *Mobiluncus* species and the characteristic odor is associated with putrescine and cadaverine that were linked to *Dialister* species. Srinivasan *et al.* [53] found that discharge is associated with agmatine and cadaverine and the odor is associated with putrescine. Agmatine, cadaverine, and putrescine were linked to 10 BV-associated bacteria including *Dialister sp. type 2* [53].

Because current understanding of BV etiology is not complete, there are no uniform criteria for diagnosing BV. In clinical practice, BV is diagnosed by Amsel criteria comprising the following four signs: 1) vaginal pH > 4.5; 2) release of "fishy" odor after adding 10% potassium hydroxide (KOH) solution to the vaginal discharge; 3) microscopy - clue cells (vaginal epithelial cells coated with bacteria); 4) white, skim milk-like vaginal discharge [54]. At least three of these four symptoms need to be present to diagnose BV. However, the criteria are not objective and rely on the perspicacity of the clinician and equipment [55]. Because of this, the misdiagnosis and recurrence rate of BV is high. In the research arena, BV is detected by the Nugent criteria that use a score of 0-10 to grade three types of bacteria in Gram-stained vaginal smears: *Lactobacillus*, *Gardnerella*, and *Mobiluncus* [56] Scores of 0-3 are considered normal, indicating that *Lactobacillus* dominates in the woman's vagina [56]. Scores of 7-10 indicate BV, showing that few *Lactobacillus* are present and *Gardnerella* and/or *Mobiluncus* dominates in the woman's vagina [56]. Scores of 4-6 are considered intermediate, indicating that *Lactobacillus, Gardnerella* and/or *Mobiluncus* are present in the woman's vagina at the same time [56]. The drawbacks of this method are as follows: 1) it is time-consuming and 2) it is not very sensitive [55-57]. For example, Nugent *et al.* [56] admitted that the Nugent

criteria have a false positive rate of 20%; recent studies have shown that only 16%–37% of women with BV diagnosed by the Nugent criteria show BV symptoms [23,58].

Both the Nugent and Amsel criteria require microscopy. Because of this, the use of these two methods is limited in some countries which do not have the necessary equipment. The BVBlue sialidase test is one of the methods to diagnose BV which does not rely on microscopy. It is a chromogenic diagnostic test - a blue or green color indicates that an elevated level of sialidase enzyme is detected in the vaginal discharge and the woman is therefore diagnosed as a BV patient [59]. It has been demonstrated that the enzyme secreted by anaerobes, such as *Bacteroides, Gardnerella, and Prevotella*, destroys mucosal tissue by hydrolyzing sialic acid from mucin [60-65]. The function of mucin is to protect the vaginal epithelium from the colonization of pathogens [66,67]. Sialidases therefore assist the adherence, colonization, and invasion of pathogens [67]. Studies have shown that the level of sialidase activity is significantly higher in BV patients compared with that in women without BV [61,62,68-70]. For example, Briselden *et al.* [61] observed elevated sialidase activity in 42 (84%) of 50 BV patients, and did not observe elevated activity in any of 19 woman without BV. The drawback of the BVBlue sialidase test is that the sensitivity is low [71]. There have been no methods to accurately diagnose BV until now.

The incomplete current understanding of the mechanism for BV makes its treatment ineffective. Metronidazole and clindamycin are commonly used antibiotics to treat BV. These two antibiotics are directed to restore the normal vaginal microbiome by killing pathogenic bacteria [40,72]. However, recurrence rates of BV in women treated with these

two antibiotics are high, with 1-month, 3-month, and 1-year recurrence rates of 23%, 43%, and 58%, respectively [72].

1.3.2 Preterm birth

Preterm birth is another major problems impacting women's reproductive health worldwide [73]. Preterm birth is birth that takes place before 37 weeks of gestational age. Every year, about 15 million babies are born preterm in the world [74]. That is more than one in ten babies [74]. The number is still rising. The societal cost for preterm birth in the United States of America alone is more than $26.2 billion in 2005 [38].

The causes of preterm birth are multifactorial. Intrauterine infection accounts for 30% of preterm birth [75]. Preterm premature rupture of membranes (PPROM) accounts for 25% of preterm birth [75]. The remaining 45% of preterm birth are idiopathic [75]. This potential mechanism which links intrauterine infection to preterm birth is that pathogens in the vagina ascend up toward the uterus, cross the placental barrier and induce preterm birth [73,75]. The potential mechanism has been supported by the previous studies that find the vaginal-related bacteria, such as *Ureaplasma urealyticum, Fusobacterium sp., Mycoplasma hominis, Streptococcus agalactiae, Peptostreptococcus sp., Staphylococcus aureus, Gardenerella vaginalis, Streptococcus viridians, Bacteroides sp., Prevotella sp., Delftia sp., Neisseria sp., Sneathia sp., and Leptotrichia sp.,* in amniotic fluid of women with preterm birth [76-80]. Among the above vaginal bacteria, *Ureaplasma urealyticum, Mycoplasma hominis, Peptostreptococcus sp., Gardenerella vaginalis, Prevotella sp.,* and *Leptotrichia sp.* were BV associated bacteria [50,51]. Some researchers thought that the

potential mechanism which links intrauterine infection to preterm birth is that bacterial products in the pathogens, such as endotoxin or peptidoglycans or lipoteichoic acid, are recognized by Toll-like receptors (TLRs) and activate nuclear factor kappa B (NF-κB), which in turn leads to the production of cytokines and the induction of preterm birth [81]. This potential mechanism has been supported by the previous studies that find the concentrations of cytokines are higher in amniotic fluid of women with preterm birth than those in women with term birth [82,83].

1.4 Metagenomic approaches

Given the important role of the vaginal microbiome in the women's reproductive health, it is important to understand what microorganisms are present in the vaginal microbiome, how many of which microbes are present in the vaginal microbiome, and what the microbes are doing. Currently, there are two popular culture-independent methods, amplicon sequencing and whole genome shotgun (WGS) sequencing, to analyze microbiomes without the need of culture [84-86]. Amplicon sequencing is the use of polymerase chain reaction (PCR) to amplify marker genes, such as the 16S ribosomal RNA (rRNA) gene, from genomic DNA and map marker genes into known sequence databases. This method can answer the questions of what microorganisms are present in the vaginal microbiome and how many of which microbes are present in the vaginal microbiome. WGS sequencing is the sequencing of genomic DNA directly from any environment (without PCR) to reveal genes present in an environment. This method can answer the questions of what microorganisms are present in the vaginal microbiome, how

many of which microbes are present in the vaginal microbiome, and what functions the microbes perform.

1.4.1 16S rRNA gene sequencing

16S rRNA genes are the most commonly used marker for microbial species identification [87]. 16S rRNA genes exist in all bacteria and contain hypervariable regions, which are species-specific and can be used to identify bacteria, and strongly conserved regions, which are located at the sides of the hypervariable regions and can be used to design primers and amplify the hypervariable regions [88]. 16S rRNA genes contain nine hypervariable regions (V1-V9) [86]. All or some of the nine regions can be used to identify bacteria. Although 16S rRNA gene sequencing is very powerful to identify bacteria, it has three drawbacks. (1) PCR amplification potentially introduces bias. Known bias can come from the design of the universal primers for 16S rRNA genes and PCR conditions for the analysis of 16S rRNA genes that can cause amplification bias and chimera formation [89,90]; (2) 16S rRNA genes are not single copy genes. The copy numbers vary from 1 up to 15 [91,92]. This leads to the underestimation or overestimation of bacterial community composition; (3) analysis of 16S rRNA genes fails to provide functional information. Although Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) can be used to predict function from 16S sequence data, PICRUSt predicts function from phylogeny based on the shared gene content and does not consider the variations among species [93]. Moreover, the prediction accuracy of PICRUSt depends on the closeness between reference genomes and query microbiome [93].

Therefore, 16S rRNA gene sequencing is not sufficient to elucidate the relationship between the vaginal microbiome and women's reproductive health, such as BV etiology.

The general workflow for 16S rRNA genes hypervariable tag sequencing technologies is (1) extraction of genomic DNA; (2) PCR amplification of the whole 16S rRNA genes or certain region of 16S rRNA genes; (3) amplicon sequencing; and (4) bioinformatics analysis, such as alignment, classification, operational taxonomic unit (OTU) analysis, and phylogenetic analysis [90,94].

1.4.2 Whole genome shotgun sequencing

An alternative to the 16S rRNA-based PCR method to identify microorganisms is WGS sequencing. WGS sequencing does not use PCR to amplify genomic DNA. Therefore, it overcomes the bias induced by PCR amplification. However, it has its own drawbacks. Genes present in an environment are sequenced based on the abundance of the organism. In general, WGS sequencing is not deep enough to detect rare organisms [86,95].

One of the important functions of WGS sequencing is to assign short reads to taxa and/or function. There are two different approaches to annotate metagenomic data, namely short reads-based annotation and assembly-based annotation. Short reads-based annotation can only provide taxonomic information. Assembly-based annotation can provide not only taxonomic information but also functional information.

1.4.2.1 Short reads-based annotation

Short reads-based annotation, as the name suggests, annotates metagenomic data directly from short reads. This method can be further classified into three groups, namely similarity-based, composition-based, and marker gene-based approach [85,96,97]. Similarity-based approach classifies metagenomic short reads based on the comparison of short reads with known sequences in a reference database [96,97]. The comparison is usually performed using Basic local alignment search tool (BLAST) [96-98]. Composition-based approach classifies metagenomic short reads based on DNA base composition signatures, such as oligonucleotide frequency [96,97]. For example, Phylopythia uses oligonucleotide composition to classify metagenomic short reads with multi-class support vector machine [99]. Marker gene-based approach classifies metagenomic short reads based on the comparison of short reads with a reference database that is built from taxonomically informative marker genes [85]. For example, Taxonomic identification and phylogenetic profiling (TIPP) uses 30 phylogenetic marker genes that are single copy and universally present in all bacteria to identify bacterial species and determine their abundance from the metagenomic short reads [100]. TIPP builds a "backbone alignment" and "backbone tree" on each of the 30 marker genes, and then uses the "backbone alignment" and "backbone tree" to find the optimal placement of the query sequence in the tree [100]. From the placement, the read can be taxonomically annotated [100].

The advantages of marker gene analysis are (1) it is a relatively rapid way to annotate bacteria from the metagenomic short reads because it compares query metagenomic short reads with marker gene database instead of all bacterial genomes [85]; (2) it is a relatively accurate way to annotate bacteria from the metagenomic short reads because marker

genes are single-copy genes [85]. It may be more accurate to determine bacterial abundance when compared with genes with multiple copies across bacterial genomes, such as 16S rRNA gene [85]. The disadvantages of marker gene analysis are (1) the choice of marker genes and the technique used to bin the reads to the markers influence the taxonomic annotation [100]; (2) all marker genes do not exist in every bacterium.

1.4.2.2 Assembly-based annotation

Assembly-based annotation, as the name suggests, annotates metagenomic data from assembly. In this method, metagenomic short reads are assembled into a single contiguous sequence (contig). Assembly-based annotation is thought to provide more accurate annotation when compared with short reads-based annotation because it works on longer sequences. Working on small number of longer sequences also reduces the computational burden when compared with working on large number of short sequences [85]. According to whether a reference database exists or not, sequence assembly could be divided into two different types of assembly, reference-based assembly and *de novo* assembly [101]. The general workflow for reference-based assembly is (1) map short reads to a reference genome; and (2) assemble short reads that are mapped to the same reference genome to contigs. The advantages of reference-based assembly are (1) contigs are longer; (2) the lengths of contigs are bigger; and (3) it does not require a lot of computational resources [101]. The disadvantages of reference-based assembly are (1) short reads that are not aligned to the reference genome are not used; and (2) contigs that are different from the reference genome cannot be produced [101]. In a *de novo* assembly, short reads that overlap are linked into contigs without the need of using a reference genome.

The advantage of *de novo* assembly is that contigs that are different from the reference genome can be produced [101]. The disadvantages of *de novo* assembly are (1) contigs are shorter; (2) the lengths of contigs are smaller; and (3) it requires a lot of computational resources [101].

Two different pipelines can be used to annotate metagenomes. The first pipeline is (1) bin contigs into different population genomes according to tetranucleotide frequencies, GC contents, and coverage of contigs [102,103]; and (2) use single-copy marker genes [103] or marker genes that are specific to a genome's inferred lineage within a reference genome tree [104] to identify the population genomes and to estimate the completeness and contamination of the population genomes. Population genomes refer to genome bins that belong to one species and some closely related strains. Binning is used because it is very difficult to get complete genomes from metagenomes at strain level. A genome bin includes a number of contigs that belong to the same species but different strains. The second pipeline is (1) predict genes present in the assembled contigs; and (2) blast the predicted genes against a reference database for taxonomic and functional annotation, such as HMP urogenital tract reference genomes, NCBI bacterial genomes, or KEGG (Kyoto Encyclopedia of Genes and Genomes) [98,105,106].

1.5 Project summary

Vaginal microbes exert a strong influence on women's reproductive health. Perturbations in the vaginal microbiome have been associated with serious sequelae, such as bacterial vaginosis (BV). Although BV has been known for centuries, the etiology of it is

multifactorial. BV is often ambiguously characterized as a shift from a *Lactobacilli*-dominated to a poly-microbial community [38]. However, recent studies found that appreciable amounts of healthy women had vaginal microbiome predominated by other various anaerobic microorganisms other than *Lactobacilli* [7], which implies that a simple understanding of the taxonomic composition of the vaginal microbiome of BV patients could not comprehensively describe the exact mechanism of BV. It has been reported that the genome of *Gardnerella vaginalis* strains that were isolated from BV patients encode for mucin degradation enzymes [52]. However, the genomes of *Gardnerella vaginalis* strains isolated from healthy women did not encode these genes [52]. The degradation of mucin facilitates colonization of pathogens and initiation of BV. Therefore, to gain a mechanistic understanding of BV, there is a need to understand what pathogens initiate it and how they do so. Amplicon sequencing can answer the question of what pathogens initiate it. Whole genome shotgun sequencing can answer both questions. However, a minority of existing research on BV used whole genome shotgun sequencing to answer the question of how pathogens initiate BV.

Sialidase has been thought to have the potential to unveil the mechanism of BV for the following reasons: (1) The enzyme can degrade mucin by hydrolyzing sialic acid from mucin [66,67]. The function of mucin is to protect the vaginal epithelium from the colonization of pathogens [66,67]; (2) Studies have shown that the level of sialidase activity is significantly higher in BV patients compared with that in women without BV [61,62,68-70]; and (3) *Prevotella*, *Gardnerella*, and *Bacteroides* species, which are reported to be associated with BV, were reported to have the ability to secrete sialidase [59,61-63,65].

Therefore, in chapter 2, we hypothesized that sialidase could serve as a biomarker for BV. The objectives of this study are to determine whether sialidase activity is elevated in Chinese BV patients, whether the relative abundance value of sialidase-encoding genes increases in Chinese BV patients, and whether the relative abundances of *Prevotella*, *Gardnerella*, and *Bacteroides* species are higher in Chinese BV patients.

Yeoman *et al.* [1] revealed that women with and without BV have distinct metabolite profiles and found metabolites that are correlated with BV symptoms. They found vaginal bacteria that were correlated with BV symptom-associated metabolites. However, they did not connect functional genes to vaginal bacteria and metabolites. BV symptom-associated metabolites are encoded by functional genes in vaginal bacteria. Therefore, studying the correlation among bacteria, functional genes, and metabolites could help better understand how pathogens initiate BV. Therefore, in chapter 3, we hypothesized that BV symptom-associated bacteria could serve as potential biomarkers to improve the development of therapeutic targets. To help test this hypothesis, we took a subset of the population from Yeoman *et al.* [1] and examined the correlation among vaginal bacteria, vaginal discharge-associated metabolites, and functional genes. We then pursued the following two objectives: (1) Identify bacteria and functional genes associated with metabolites that Yeoman *et al.* [1] found to be correlated with vaginal discharge, and (2) Identify which identified bacteria were correlated with the identified functional genes.

The vaginal microbiomes are mainly dominated by *Lactobacilli* during pregnancy [10,13-15]. The predominance of *Lactobacilli* in the vagina has been reported to be related to high

estrogen levels and glycogen content in the vaginal epithelium [6,11]. Estrogen levels increase gradually during pregnancy, reaching a maximum near term, and drop down quickly during labor [107]. Glycogen is converted to lactic acid mainly by *Lactobacilli*. Therefore, in chapter 4, we hypothesized that the abundance of *Lactobacilli* and functional genes for metabolic activity in *Lactobacilli* would increase as pregnancy progresses and decrease following parturition. The objectives of this study are to determine whether the abundance of *Lactobacilli* and functional genes for metabolic activity in *Lactobacilli* increase as pregnancy progresses and decrease following parturition.

CHAPTER 2. TAXONOMIC AND FUNCTIONAL GENE COMPOSITION DIFFERS BETWEEN CHINESE WOMEN WITH AND WITHOUT BACTERIAL VAGINOSIS

2.1 Abstract

Bacterial vaginosis (BV) is common vaginal disorder globally in women of reproductive age, and has been associated with serious sequelae such as sexually transmitted diseases and preterm birth. Although BV has been known for centuries, the etiology of it is still poorly understood. Sialidase has been thought to have the potential to unveil the mechanism of BV. It assists the adherence, colonization, and invasion of pathogens. This study was conducted to determine whether sialidase could serve as a biomarker for BV. We used the BVBlue sialidase test to measure the sialidase activity, 454-pyrosequencing of amplicons from the V1-V3 region of the 16S rRNA gene to classify sequence to bacteria, and Illumina high-throughput sequencing to classify sequence to bacteria and functional genes in microbial samples from 38 Chinese women with and without BV from 20 to 52 years (means ± SEM 35.6 ± 1.2 years). The samples were collected by swabs from cervix, middle, and upper vagina walls. Elevated sialidase activity (>= 7.8U) was detected in all BV patients but not in any women without BV. Analysis of 16S rRNA gene and whole genome shotgun sequences data revealed that *Prevotella* and *Gardnerella* species as well as thiol-activated cytolysin (K11031) and putative peptidoglycan lipid II flippase (K03980) were more abundant in BV patients. We conclude that the sialidase test could separate diseased vaginal microbiomes from non-diseased ones. *Prevotella* and *Gardnerella* species were associated with high sialidase

levels. *Gardnerella vaginalis* may initiate BV through the produce of thiol-activated cytolysin.

2.2 Introduction

Bacterial vaginosis (BV) is one of the most common, yet major problems impacting women's reproductive health worldwide [22]. The prevalence of BV in the United States is about 29.2% or approximately 21.2 million women of reproductive age [23] and is associated with serious sequelae. In non-pregnant women, BV has been linked predominantly to the upper genital tract and poses an increased risk for sexually transmitted infections [3,24-27] and susceptibility to human immunodeficiency virus (HIV) [27-29]. In pregnant women, BV has been linked to miscarriages [27,30-32], post-abortion infections [27,33], postpartum endometritis [27,34,35], preterm pre-labor rupture of membranes (PPROM) [36], and preterm birth [3,27,37,38]. The United States annually spends nearly $1 billion on treating BV-related pregnancy complications [39].

Considering medical and economic importance of vaginal health, there is a great interest in understanding of the etiology of BV. BV is often ambiguously characterized as perturbations in the normal vaginal microbiome [7], and is believed to be a result of a shift from a *Lactobacilli*-dominated to a poly-microbial community [38]. For example, BV patients are believed to have less *Lactobacilli* and more *Gardnerella* and *Mobiluncus* compared to healthy asymptomatic women. However, species like *Gardnerella vaginalis* and *Mobiluncus curtisii*, which have been traditionally associated with BV, are also commonly found in healthy asymptomatic women [49]. Moreover, recent studies suggest that *Lactobacilli* do not always predominate in the vaginal microbiome. In 10 to 42% of healthy asymptomatic women, *Lactobacilli* were outnumbered by other vaginal microbes, such as *Atopobium vaginae, Leptotrichia and Megasphaera*, which have the ability to

produce lactic acid and keep the vaginal environment acidic [9,41-48]. This implies that a simple understanding of the taxonomic composition of the vaginal microbiome of BV patients could not comprehensively describe the exact mechanism of BV.

Sialidase (or neuraminidase) production is thought to play a mechanistic role in the development of BV. It has been demonstrated that sialidase secreted by vaginal bacteria degrades mucin by hydrolyzing sialic acid from mucin [66,67]. The function of mucin is to protect the vaginal epithelium from the colonization of pathogens [66,67]. Sialidases therefore assist the adherence, colonization, and invasion of pathogens [67]. Studies have shown that the level of sialidase activity is significantly higher in BV patients compared with that in women without BV [61,62,68-70]. For example, Briselden *et al.* [61] observed elevated sialidase activity in 42 (84%) of 50 BV patients, and did not observe elevated activity in any of 19 woman without BV. Additionally, *Prevotella*, *Gardnerella*, and *Bacteroides* species, which are reported to be associated with BV, were reported to have the ability to secrete sialidase [59,61-63,65].

Therefore, we hypothesized that sialidase could serve as a biomarker for BV. The objectives of this study are to determine whether sialidase activity is elevated in Chinese BV patients, whether the relative abundance value of sialidase-encoding genes increases in Chinese BV patients, and whether the relative abundances of *Prevotella*, *Gardnerella*, and *Bacteroides* species are higher in Chinese BV patients. We used the BVBlue sialidase test to measure the sialidase activity and analyzed 16S rRNA genes and whole genome shotgun sequences in microbial samples from Chinese women with and without

BV to identify bacterial species and functional genes that were more abundant in Chinese BV patients.

2.3 Methods

2.3.1. Ethics statement

This study was approved by NanFang Hospital in Guangzhou, China (BV_2011_01P). Informed written consent was obtained from all study participants prior to sample collection.

2.3.2. Sample collection

For this study, vaginal samples were collected from 38 Chinese women. None of the women were pregnant, and all were of reproductive age from 20 to 52 years (means ± SEM 35.6 ± 1.2 years). All subjects lived in Guangzhou, China. The sample descriptions are given in Table 1. Twenty-six patients were diagnosed for BV by a Chinese physician who collected their vaginal samples by taking swabs from cervix, middle, and upper vagina walls. Two swabs were taken from each woman. One swab was used for the BVBlue sialidase test kit (Huajin Biotechnology Co. Ltd., China, Cat. No. 2400018) that is a chromogenic diagnostic test; a blue or green color indicates an elevated level of sialidase enzyme in the vaginal discharge [59]. The BVBlue sialidase test kit detects vaginal fluid sialidase activity at levels of >=7.8 U [59]. One unit of sialidase activity is defined as the amount of enzyme required to liberate 1 nmol of substrate/ml/min at 37°C [59]. The second swab was saved in a sterile cryovial with 1 ml saline for the purpose of DNA extraction and further analysis.

2.3.3. Genomic DNA extraction

The vaginal swabs were vigorously agitated to dislodge the bacterial cells. The cells were pelleted by centrifugation at 5,000 g for 10 min. Bacterial genomic DNA was extracted using QIAamp DNeasy Blood and Tissue Kit (QIAGEN, Hilden, Germany; Cat. No. 69504) following the manufacturer's instructions. Briefly, the bacterial pellet was re-suspended in 200 µl of saline with 20 µl of proteinase K solution and 200 µl of lysis buffer (Buffer AL). The mixtures were homogenized by vortexing, and incubated at 56°C for 30 min. After incubation, 200 µl of 100% ethanol was added and mixed by vortexing. The mixture was transferred into QIAamp spin column and centrifuged at 8,000 g for 1 min. The QIAamp spin column was replaced in a new 2 ml collection tube and washed twice with 500 µl buffer AW1 and AW2 provided in the kit. The genomic DNA was eluted with 100 µl of elution buffer (Buffer AE). Concentration of DNA was determined using NanoDrop ND-1000 spectrophotometer (Thermo Electron Corporation, USA). The integrity and fragment size of the extracted DNA were determined by 1% agarose gel electrophoresis.

2.3.4. Pyrosequencing

The V1-V3 region of the 16S ribosomal RNA gene of 38 samples was amplified from extracted DNA by polymerase chain reaction (25 cycles of 94°C (30 s), 48°C (30 s), 72°C (2 min)) using primers 27f-YM (CGTATCGCCTCCCTCGCGCCATCAG-AGAGTTTGATYMTGGCTCAG); and 534r (CTATGCGCCTTGCCAGCCCGCTCAG-[MID tag 1 – 50]-

ATTACCGCGGCTGCTGGCA). The primers were validated previously and shown to broadly encompass vaginal bacteria [108]. Negative control PCRs were performed without 1) template DNA and 2) without polymerase. No products were detected in any negative control. The amplicons were pyrosequenced on 454 FLX-Titanium Technology at the J. Craig Venter Institute (Rockville, MD).

2.3.5. Illumina high-throughput sequencing

Four out of 38 DNA samples (IDs: SMUB9, SMUB27, SMUN23, SMUN25) were selected for metagenomic analysis by sequencing on Illumina HiSeq 2000. The four purified DNA samples were solubilized in EB (1 mM TrisHCl pH 8), amplified by multiple displacement amplification (MDA) using phi29 enzyme mix (GenomiPhi V2 DNA Amplification Kit; GE Healthcare), according to the manufacturer's protocol. 10-50 ng of each sample was used as a template and reactions were carried out for 16 hours. 500 ng of amplified DNA was used for libraries construction with standard Illumina protocols with few modifications. DNA was sheared using the Covaris S2 or E210 system, followed by ligation of sequencing adaptors containing unique barcode sequences to allow pooling of all samples. Libraries were purified using Agencourt AMPure XP beads, and DNA concentration was measured using the Agilent High Sensitivity DNA Kit. Quality control steps were completed with qPCR using the KAPA Biosystems Library Quantification Kit. Pooled libraries were sequenced on one lane of a HighSeq 2000 machine, paired-end 2x101bp.

2.3.6. 16S rRNA sequencing data analysis and statistics

The resulting sequences were quality trimmed using the methods described in a previous study [109], with the minor modification of the use of an average quality score ≤ 30 instead of 25. A total of 553,888 high quality sequence reads were generated from 38 samples, with an average sequence length 484 nt (Table 2). For taxonomic analysis, the resulting sequences were assigned to operational taxonomic units (OTUs) as previously described [1,109,110]. Briefly, sequences were aligned against the Silva database and pre-clustered using mothur [111]. The remaining sequences were clustered at 97% sequence similarity using ModalClust, utilizing a complete linkage clustering method (https://bitbucket.org/msipos/modalclust). OTUs detected in fewer than three samples and three times were removed as possible artifacts. The taxonomic similarities among microbial communities were compared using Bray-Curtis dissimilarity statistics. To calculate relative abundance of each OTU in each sample, the number of sequences in each OTU in the OTU table was divided by the total number of sequences in each sample. The relative abundance matrix was transformed by square root to reduce the effect of more abundant over less abundant OTUs. Resemblance matrices and principal coordinate analysis (PCoA) plots were created and visualized in Primer (Primer-E 2007). To test the null hypothesis that there was no difference between Chinese women with and without BV, we used permutational multivariate analysis of variance (PERMANOVA) in Primer (Primer-E 2007). The representative sequence in each OTU was assigned to genera using RDP classifier v2.5 with a bootstrap cutoff of 80% [112]. Significantly different bacterial genera were identified using linear discriminant analysis (LDA) effect size (LEfSe) Galaxy version with default parameters [113] and visualized with a heatmap which was generated using heatmap.2 function of the gplots package in R. Women with

and without BV were two classes without any subclasses. Shannon diversity indices were calculated in mothur [111]. Significantly different Shannon diversity indices were detected using a Kruskal-Wallis rank sum test [114]. P-values < 0.05 were regarded as significant unless specifically stated otherwise.

2.3.7. Metagenomic sequencing data analysis and statistics

2.3.7.1 Reads trimming

The qualities of the raw Illumina sequencing data were first assessed using FastQC v0.10.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The first few bases of each sequence, which showed a bias in the FastQC report, were trimmed using the FASTX toolkit [115]. Leading and trailing low quality bases (Q score below 20) and sequences shorter than 50 nucleotides and with an average quality score < 20 were removed using trimmomatic v0.30 [116]. This quality filtering resulted in the quality values of all bases being higher than 28 in the FastQC report. To remove potential contamination with human DNA, the remaining reads were further mapped against human genome reference sequence (NCBI build 37 p13) with one mismatch permitted per read in Bowtie 2 [117].

2.3.7.2 Assembly #1

To obtain a reliable assembly, the resulting clean reads were aligned against the HMP urogenital tract reference genomes downloaded on January 10, 2015 (Table 3). The read alignments with mapping qualities > 20 were considered to be true alignments. The reads whose breadth of coverage (defined as the percentage of bases covered in the entire

reference genome) was more than 2% at least one in sample were further assembled to contigs using SOAPdenovo v2.04, with k-mer corresponding to the longest N50 contig length (Table 4) [118]. The assembled contigs of each sample were aligned against HMP urogenital tract reference genomes using ABACAS with default parameters, except 80% minimum percent identity [119]. The contigs that shared at least 80% identity with the HMP urogenital tract reference genomes were called “assembly #1” and were combined with the following “assembly #2” and “assembly #3” for functional gene annotation.

2.3.7.3 Assembly #2 and assembly #3

The reads that were not used in “assembly #1” were assembled into contigs as previously described [103]. Briefly, the paired-end reads were digital normalized to k-mer coverage (defined as the number of times a k-mer was found in a read) of 20 using normalize-by-median.py in khmer v1.3. [120]. Highly repetitive k-mers, which could belong to multiple species, were removed using filter-below-abund.py with C=100 in khmer v1.3. [120]. The resulting paired-end reads were *de novo* assembled using Ray Meta at k=23 [121]. Scaffold coverage and length were calculated using Bowtie 2 [117] and samtools v0.1.18 [122]. Tetranucleotide frequency, GC content, and conserved marker genes were identified as described [103]. The binning of scaffolds into genomes was performed as described [103]. Scaffolds whose sequencing depths were bigger than 20X in any two subjects and which were classified by essential single copy gene were chosen (Table 5). The unassembled reads associated with these scaffolds were extracted and reassembled using SOAPdenovo v2.04 at k=23, thus, resulting new assembly [118]. The new assemblies were called “assembly #2”. The unassembled reads that were not used in “assembly #1” or “assembly

#2" were *de novo* assembled using Ray Meta at k=23 [121]. This assembly was called "assembly #3". "Assembly #2" and "assembly #3" were combined with "assembly #1" for the taxonomic and functional annotation of genes.

2.3.7.4 Taxonomic and functional annotation of genes

Assemblies 1, 2, and 3 were combined for each sample for the taxonomic and functional annotation of genes. A total of 13 Mbp of contigs (>= 300 bp) were generated in the combined contigs from four samples with longest contig length of 36 Kbp (Table 6). Microbial genes were also searched for in all contig sequences longer than 300 bp using the Prokaryotic Dynamic Programming Genefinding Algorithm (Prodigal) [123]. The predicted genes of all samples were combined to create a gene set. To build a non-redundant gene set, genes with 95% sequence identity and aligned length covered over 90% of the sequence were clustered together using cd-hit v4.6.1 [124]. To create a gene abundance matrix in which both pairs that mapped to the same gene would only be counted once, clean short reads in each sample were mapped back to the non-redundant gene set using bwa v0.7.5a [125] and samtools v0.1.18 [122]. For taxonomic annotation, non-redundant genes were queried using BLASTN [98] (default parameters except E=1e-5) against the HMP urogenital tract reference database downloaded on January 10, 2015 (Table 3) and the NCBI complete bacteria genome, downloaded on September 6, 2014. For functional gene annotation, non-redundant genes were queried using DIAMOND [105] (default parameters except E=1e-5) against the KEGG database (v56) [106]. Only the best hits with at least 80% identity over 100 bp-aligned lengths were selected for comparison between Chinese women with and without BV.

2.3.7.5 Genomic analyses

The combined contigs from each sample (>= 300 bp) were aligned to the HMP urogenital tract reference genomes downloaded on January 10, 2015 (Table 3) using BLASTN [98] (default parameters except E=1e-5). Only the best hits with at least 80% identity over 100 bp-aligned lengths were selected. Contigs aligned against the *Gardnerella vaginalis 315-A, Gardnerella vaginalis 409-05, Gardnerella vaginalis ATCC 14019, Gardnerella vaginalis HMP9231*, and *Prevotella timonensis CRIS 5C-B1* reference genome in each sample were compared with the corresponding reference genome using GView server with BLAST atlas analysis type and default parameters except e-value < 0.00001 [126].

2.4 Results

2.4.1 Sialidase enzyme analysis

To determine whether sialidase activity is elevated in Chinese BV patients but not in Chinese women without BV, we used the BVBlue sialidase test to measure the sialidase activity in samples from Chinese women with and without BV. A blue or green color was observed in all samples collected from Chinese BV patients indicating elevated sialidase activity was detected in all patients with BV. A yellow color was observed in all samples collected from Chinese women without BV indicating sialidase activity was not observed to increase in any patients without BV. Therefore, the test was able to distinguish between women with and without BV.

2.4.2 Diversity and microbial community structure analyses

To further determine whether the sialidase test could separate women with and without BV, we examined the microbiome diversity and microbial community structure of women with and without BV. The vaginal microbiomes of BV patients were more diverse than those of women without BV (Figure 1). Bray-Curtis dissimilarity-based PCoA plot exhibited a disparate clustering between women with and without BV (PERMANOVA p=0.001; Figure 2). Therefore, women with and without BV harbored distinctly different vaginal microbiomes.

2.4.3 Functional analysis of metagenomic sequence data

Elevated sialidase enzyme were detected in all Chinese BV patients and not detected in any Chinese women without BV. It implies that sialidase-encoding genes are more abundant in Chinese BV patients when compared with Chinese women without BV. To determine possible genetic underpinnings for this measurement, we annotated all open reading frames to KEGG orthologous groups. Gene sequence encoding a sialidase enzyme was not found in all Chinese women. It means that the sequencing of all Chinese women may be not deep to find sialidase enzyme. To find an alternative to sialidase-encoding genes, we compared all predicted enzymes of Chinese women with BV with those of Chinese women without BV. We found that the relative abundance of 223 KEGG orthologous groups were higher in the vaginal microbiomes of both Chinese BV patients when compared with both Chinese women without BV (Table 7). Among these 223 KEGG orthologous groups, thiol-activated cytolysin (K11031), which is a toxin, and putative peptidoglycan lipid II flippase (K03980), which is virulence factor, may serve as biomarkers for BV (Figure 3).

2.4.4 Taxonomic analysis of 16S rRNA gene and whole genome shotgun sequence data

To determine whether the relative abundance value of *Prevotella*, *Gardnerella*, and *Bacteroides* species were higher in Chinese BV patients when compared with Chinese women without BV, the LEfSe was used to identify which bacterial genera were more abundant in BV patients [113]. At a p-value of 0.05, 16S rRNA sequence data revealed that eight bacterial genera including *Prevotella* and *Gardnerella* were significantly more abundant in Chinese BV patients (Figure 4; Table 8). Genus *Bacteroides* was not found to be significantly more abundant in Chinese BV patients (Figure 4; Table 8). At the strain level, metagenomic sequence data revealed that the relative abundance of *Gardnerella vaginalis 315-A*, *Gardnerella vaginalis 409-05*, *Gardnerella vaginalis ATCC 14019*, *Gardnerella vaginalis HMP9231*, and *Prevotella timonensis CRIS 5C-B1* were higher in the vaginal microbiomes of both Chinese BV patients when compared with both Chinese women without BV (Figure 5; Table 9). The relative abundance of bacterial strains affiliated with genus *Bacteroides* were not found to be higher in the vaginal microbiomes of both Chinese BV patients when compared with both Chinese women without BV (Table 9).

2.4.5 Genomic analysis

To find out how many nucleotides of each of the *Gardnerella vaginalis 315-A*, *Gardnerella vaginalis 409-05*, *Gardnerella vaginalis ATCC 14019*, *Gardnerella vaginalis HMP9231*, and *Prevotella timonensis CRIS 5C-B1* reference genome was

covered in our assemblies, we aligned individual assembly against all HMP urogenital tract reference genomes and compared contigs that were aligned to the above five reference genomes with the corresponding reference genome. More nucleotides from the *Gardnerella vaginalis 315-A, Gardnerella vaginalis ATCC 14019, Gardnerella vaginalis HMP9231,* and *Prevotella timonensis CRIS 5C-B1* reference genomes had the alignment with the assemblies from two Chinese BV patients when compared with two Chinese women without BV (Figure 6).

2.5 Discussion

Our study revealed that elevated sialidase activity was detected in all BV patients but not in any women without BV and women with and without BV harbored distinctly different vaginal microbiomes. These suggest that sialidase test could separate diseased vaginal microbiomes from non-diseased ones. The separation of diseased vaginal microbiome from non-diseased one was also observed in the previous study where researchers used Amsel and Nugent criteria to diagnose BV and revealed that microbiome richness and diversity are increased in BV patients and the abundances of vaginal bacterial species in BV patients significantly differ from those in healthy controls [127].

In our study, we observed that *Prevotella* and *Gardnerella* species were more abundant in women with BV. These findings support the findings from previous studies where researchers reported that *Prevotella*, *Gardnerella*, and *Bacteroides* species have the ability to secrete sialidases [59,61-63,65]. Our whole genome shotgun sequences data further revealed that *Gardnerella vaginalis 315-A, Gardnerella vaginalis 409-05, Gardnerella*

vaginalis ATCC 14019, *Gardnerella vaginalis HMP9231*, and *Prevotella timonensis CRIS 5C-B1* were more abundant in both BV patients. Sialidases-encoding genes were found in the *Gardnerella vaginalis 315-A* and *Gardnerella vaginalis ATCC 14019* reference genome. This implies that sialidases may be produced by *Gardnerella vaginalis 315-A* and *Gardnerella vaginalis ATCC 14019*.

Sialidases-encoding genes were not found in all samples. This might be because the coverage of our metagenome data sets was low. However, we found that the relative abundance of thiol-activated cytolysin (K11031) was higher in BV patients than women without BV. Thiol-activated cytolysin is a toxin, and is reported to be made by *Gardnerella vaginalis* [48,128]. Thiol-activated cytolysin plays a role in BV pathogenesis through lysing vaginal epithelial cells and activating the generation of IgA antibodies, further inducing mucosal immune response [48]. This implies that *Gardnerella vaginalis* may also initiate BV through the produce of thiol-activated cytolysin.

One of the limitations of this study is the small sample size. To alleviate the influence of small sample size, only bacterial strains or KEGG orthologous groups whose relative abundance values were higher in all individuals in one group than the other group, were considered to be more abundant ones. The second limitation of this study is low metagenome coverage. Studies based on 16S rRNA gene sequences revealed that there were about 400 bacteria in the vagina [9,10,12]. We assume the average bacterial genome size is 3MB. The rough estimate of the number of bases in the vaginal metagenome at 30 X sequencing depth was 36,000 MB (3 Mb * 400 * 30). The highest number of bases in

our quality-filtered metagenome was about 568 MB, which corresponded to 2% of metagenome coverage. To alleviate the influence of low metagenome coverage, we used relative abundance values and only consider high abundance bacterial strains or KEGG orthologous groups.

2.6 Conclusions

Sialidase test could separate diseased vaginal microbiomes from non-diseased ones. *Prevotella* and *Gardnerella* species were associated with high sialidase levels. *Gardnerella vaginalis* may initiate BV through the produce of thiol-activated cytolysin.

CHAPTER 3. IDENTIFICATION OF VAGINAL DISCHARGE ASSOCIATED BACTERIA BY METAGENOMIC AND NETWORK ANALYSIS

3.1 Abstract

Bacterial vaginosis (BV) is a common vaginal disorder globally in women of reproductive age, and has been associated with serious sequelae such as sexually transmitted diseases and preterm birth. Although BV has been known for centuries, the etiology of it is still poorly understood. Because of this, the diagnosis and treatment of BV are ineffective. Yeoman *et al.* [1] found that thirty-three metabolites correlated with the presence of BV-associated vaginal discharge. This study was conducted to identify discharge-associated bacteria that can serve as potential biomarkers to improve the development of therapeutic targets. To better understand how pathogens are involved in the production of BV symptoms, we took a subset of the population (n=4) from Yeoman *et al.* [1]. All women were of reproductive age from 20 to 40 years. The samples were collected by swabs, scrapings, and lavage. 454-pyrosequencing of amplicons from the V1-V3 region of the 16S rRNA gene was used to classify sequence to bacteria. Illumina high-throughput sequencing was used to classify sequence to bacteria and functional genes. We then used Spearman's rank correlation coefficients to examine the correlation among vaginal bacteria, vaginal discharge-associated metabolites, and functional genes. Sixty-five functional genes and sixteen bacterial genera were significantly linked to discharge-associated metabolites (r=1, p=0). Twelve bacterial genera were significantly positively linked to eleven functional genes (r=1, p=0). Among the twelve bacterial genera and eleven functional genes, the genera *Aquabacterium*, *Chryseobacterium*,

Elizabethkingia, *Finegoldia*, *Morganella*, *Negativicoccus*, *Phenylobacterium*, *Propionibacterium*, *Serratia*, *Siphonobacter*, and the enzymes 'malate dehydrogenase' (K00024), '5-amino-6-(5-phosphoribosylamino) uracil reductase' (K11752), and 'undecaprenyl-diphosphatase' (K06153) were positively associated with 2-Methyl-2-hydroxybutanoic acid (r=1, p=0) that was the only metabolite revealed by Yeoman *et al.* [1] that positively correlated with discharge. Moreover, these three KEGG orthologous groups were positively linked to these ten bacterial genera (r=1, p=0). We conclude that the bacteria *Aquabacterium*, *Chryseobacterium*, *Elizabethkingia*, *Finegoldia*, *Morganella*, *Negativicoccus*, *Phenylobacterium*, *Propionibacterium*, *Serratia*, and *Siphonobacter* are potential therapeutic targets.

3.2 Introduction

Bacterial vaginosis (BV) is one of the most common, yet major problems impacting women's reproductive health worldwide [22]. The prevalence of BV in the United States is about 29.2% or approximately 21.2 million women of reproductive age [23] and is associated with serious sequelae. In non-pregnant women, BV has been linked predominantly to the upper genital tract and poses an increased risk for sexually transmitted infections [3,24-27] and susceptibility to human immunodeficiency virus (HIV) [27-29]. In pregnant women, BV has been linked to miscarriages [27,30-32], post-abortion infections [27,33], postpartum endometritis [27,34,35], preterm pre-labor rupture of membranes (PPROM) [36], and preterm birth [3,27,37,38]. The United States annually spends nearly $1 billion on treating BV-related pregnancy complications [39].

The etiology of BV is multifactorial. BV is often ambiguously characterized as perturbations in normal vaginal microbiome [7], and believed is a result of a shift from a *Lactobacilli*-dominated to a poly-microbial community [38]. For example, BV patients are believed to have less *Lactobacilli* and more *Gardnerella* and *Mobiluncus* compared with healthy asymptomatic women. The potential mechanisms supporting the ambiguous characterization of BV are: (1) *Lactobacillus* species produce lactic acid to maintain an acidic vaginal environment (pH < 4.5), hydrogen peroxide and bacteriocin to kill pathogens in vagina; and (2) anaerobic bacteria produce polyamines that cause a "fishy" odor in the vaginal discharge, as well as mucinases and sialidases enzymes, which destroy mucosal tissue and facilitate colonization of pathogens in the vagina [40]. However, recent studies suggest that *Lactobacilli* do not always predominate in the vaginal

microbiomes of healthy women. In 10 to 42% of healthy asymptomatic women, *Lactobacilli* were outnumbered by other vaginal microbes, such as *Atopobium vaginae*, *Leptotrichia*, and *Megasphaera* [9,41-48]. Species like *Gardnerella vaginalis* and *Mobiluncus curtisii*, which have been traditionally associated with BV, are also commonly found in healthy asymptomatic women [49]. The incomplete current understanding of the mechanism for BV makes its diagnosis and treatment ineffective.

Gradually, researchers found that a simple understanding of the taxonomic composition of the vaginal microbiomes of BV patients could not comprehensively describe the exact mechanism of BV. For example, it has been reported that the genome of *Gardnerella vaginalis* strains that were isolated from BV patients encode for mucin degradation enzymes [52]. However, the genomes of *Gardnerella vaginalis* strains isolated from healthy women did not encode these genes [52]. The degradation of mucin facilitates colonization of pathogens and initiation of BV. Therefore, to gain a mechanistic understanding of BV, there is a need to understand what pathogens initiate it and how they do so. However, a minority of existing research on BV focused on answering the question of how pathogens initiate BV.

To answer the question of how pathogens initiate BV, Yeoman *et al.* [1] and Srinivasan *et al.* [53] related vaginal bacteria to distinct BV metabolites. Specifically, Yeoman *et al.* [1] revealed that women with and without BV have distinct metabolite profiles and found metabolites that are correlated with BV symptoms: white, skim milk-like vaginal discharge; “fishy” vaginal odor; vaginal itching; and pain during urination. Yeoman and

his colleagues [1] related *Mobiluncus* species to vaginal discharge via 2-methyl-2-hydroxybutanoic acid; related *Dialister* species to vaginal odor via putrescine and cadaverine; and related *Gardnerella* species to pain via diethylene glycol. Similarly, Srinivasan *et al.* [53] also revealed that metabolite profile of BV patients differ from that of women without BV. However, among the metabolites that distinguish metabolite profiles of BV patients from those of women without BV, only 12 metabolites were found in both studies [1,53]. Srinivasan and his colleagues [53] also found BV symptom-associated metabolites [53]. They found that discharge is associated with agmatine and cadaverine and the odor is associated with putrescine [53]. Agmatine, cadaverine, and putrescine were linked to ten BV-associated bacteria including *Dialister sp. type 2* [53].

Although Yeoman *et al.* [1] and Srinivasan *et al.* [53] found vaginal bacteria that were correlated with BV symptom-associated metabolites, neither study connected functional genes to vaginal bacteria and metabolites. BV symptom-associated metabolites are encoded by functional genes in vaginal bacteria. Therefore, studying the correlation among bacteria, functional genes, and metabolites could help better understand how pathogens initiate BV.

Therefore, we hypothesized that BV symptom-associated bacteria could serve as potential biomarkers to improve the development of therapeutic targets. To help test this hypothesis, we took a subset of the population from Yeoman *et al.* [1] and examined the correlation among vaginal bacteria, vaginal discharge-associated metabolites, and functional genes. We then pursued the following two objectives: (1) Identify bacteria and

functional genes associated with metabolites that Yeoman *et al.* [1] found to be correlated with vaginal discharge, and (2) Identify which identified bacteria were correlated with the identified functional genes. We followed metagenomic approaches using samples from women with and without BV from the US to determine the taxonomic and functional composition of the vaginal microbiome and build a network among vaginal bacteria, functional genes, and metabolites.

3.3 Methods

3.3.1. Ethics statement

This study was approved by the Institutional Review Board of the University of Illinois at Urbana-Champaign, USA (#05079, obtained 10-31-11) and Carle Foundation Hospital in Urbana, USA (#89689-4, obtained 06-17-11). Informed written consent was obtained from all study participants prior to sample collection.

3.3.2. Samples collection

For this study, vaginal samples were collected from five US women, a subset of the population from Yeoman *et al.* [1]. None of the women were pregnant, and all were of reproductive age from 20 to 40 years. All subjects lived in Champaign-Urbana, Illinois, USA, and represented three ethnicities (Caucasian, Asian American, and African American). The sample descriptions are given in Table 10. The patients were evaluated for BV via the Amsel criteria (vaginal discharge, pH > 4.5, odor, clue cells) [54] by a trained clinician with >20 years experience. The samples were collected by swabs, scrapings, and lavage. Swabs were collected from the outer surface of the cervix, the

large posterior portion of the fornix, and the outer third of the vaginal canal. Scrapings were collected from the upper third and lower third of the vagina. Ectocervicovaginal lavage samples were collected by injecting 15ml of sterile saline solution (Health Care Logistics, Circleville, Ohio) in a continuous stream toward the cervix and then collected in the posterior fornix using a syringe fitted with a catheter. The scheme of sample collection sites is described in a previous study [129]. Each sample was aspirated five times prior to removal and storage. All samples were stored in sterile saline solution (Health Care Logistics, Circleville, Ohio) in 15-ml tubes, frozen immediately upon collection, and stored at -80°C for up to three years until use. Vaginal smears were also obtained for each subject at the time of clinical examination by rolling a swab across and along the length of the vaginal wall and then onto a glass slide, which was stored at -80°C until being Gram stained for independent evaluation by a second trained professional, also with >20 years experience, according to the Nugent criteria [56].

3.3.3. Genomic DNA extraction

Genomic DNA was isolated from 0.5 ml aliquots of each collected sample. To each aliquot, a solution of 125 µl of 0.5 M Na.EDTA (pH 8) containing 75 mg/ml of lysozyme was added and samples were lysed at 37°C for 30 min. To each sample, 70 µl of sodium-dodecyl-sulfate was added and the protein content was degraded by incubation with proteinase K (0.05 mg/ml final concentration) at 55°C for 30 min. Further lysis was facilitated by three cycles of rapid freeze-thaw involving immersion in dry-ice/ethanol slurry until frozen followed by heating in a heatblock to 37°C. Protein was precipitated by the addition of 70 µl of 5 M NaCl and 30 min incubation on ice. Precipitated protein

was removed by centrifugation at 14,000 rpm for 20 min and residual organic matter was removed by phenol:chloroform washes followed by ethanol precipitation to remove residual salts.

3.3.4. Pyrosequencing

The V1-V3 region of the 16S ribosomal RNA gene of five samples was amplified from extracted DNA by polymerase chain reaction (25 cycles of 94°C (30 s), 48°C (30 s), 72°C (2 min)) using primers 27f-YM (CGTATCGCCTCCCTCGCGCCATCAG-AGAGTTTGATYMTGGCTCAG); and 534r (CTATGCGCCTTGCCAGCCCGCTCAG-[MID tag 1 – 50]-ATTACCGCGGCTGCTGGCA). The primers were validated previously and shown to broadly encompass vaginal bacteria [108]. Negative control PCRs were performed without 1) template DNA and 2) without polymerase. No products were detected in any negative control. The amplicons were pyrosequenced on 454 FLX-Titanium Technology at the J. Craig Venter Institute (Rockville, MD).

3.3.5. Illumina high-throughput sequencing

The purified DNA samples collected by lavage were solubilized in EB (1 mM TrisHCl pH 8), amplified by multiple displacement amplification (MDA) using phi29 enzyme mix (GenomiPhi V2 DNA Amplification Kit; GE Healthcare), according to the manufacturer's protocol. 10-50 ng of each sample was used as a template and reactions were carried out for 16 hours. 500 ng of amplified DNA was used for libraries construction with standard Illumina protocols with few modifications. DNA was sheared

using Covaris S2 or E210 system, followed by ligation of sequencing adaptors containing unique barcode sequences to allow pooling of all samples. Libraries were purified using Agencourt AMPure XP beads, and DNA concentration was measured using the Agilent High Sensitivity DNA Kit. Quality control steps were completed with qPCR using the KAPA Biosystems Library Quantification Kit. Pooled libraries were sequenced on one lane of HighSeq 2000 machine, paired-end 2x101bp.

3.3.6. 16S rRNA sequencing data analysis and statistics

The resulting sequences collected by swabs, scrapings, and lavage were combined by subject. The combined sequences were quality trimmed using the methods described in a previous study [109], with the minor modification of the use of an average quality score ≤ 30 instead of 25. A total of 437,176 high quality sequence reads were generated from five samples, with an average sequence length 455 nt (Table 11). For taxonomic analysis, the resulting sequences were assigned to operational taxonomic units (OTUs) as previously described [1,109,110]. Briefly, sequences were aligned against the Silva database and pre-clustered using mothur [111]. The remaining sequences were clustered at 97% sequence similarity using ModalClust, utilizing a complete linkage clustering method (https://bitbucket.org/msipos/modalclust). The representative sequence in each OTU was assigned to genera using RDP classifier v2.5 with a bootstrap cutoff of 80% [112]. To calculate relative abundance of each genus in each sample, the number of sequences in each genus in the genus table was divided by the total number of sequences in each sample.

3.3.7. Metagenomic sequencing data analysis and statistics

3.3.7.1 Reads trimming

The qualities of the raw Illumina sequencing data were first assessed using FastQC v0.10.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The first few bases of each sequence, which showed a bias in the FastQC report, were trimmed using the FASTX toolkit [115]. Leading and trailing low quality bases (Q score below 20) and sequences shorter than 50 nucleotides and with an average quality score < 20 were removed using trimmomatic v0.30 [116]. This quality filtering resulted in the quality values of all bases being higher than 28 in the FastQC report. To remove potential contamination with human DNA, the remaining reads were further mapped against human genome reference sequence (NCBI build 37 p13) with one mismatch permitted per read in Bowtie 2 [117].

3.3.7.2 Assembly #1

To obtain a reliable assembly, the resulting clean reads were aligned against the HMP urogenital tract reference genomes downloaded on January 10, 2015 (Table 3). The read alignments with mapping qualities > 20 were considered to be true alignments. The reads whose breadth of coverage (defined as the percentage of bases covered in the entire reference genome) was more than 2% at least one in sample were further assembled to contigs using SOAPdenovo v2.04, with k-mer corresponding to the longest N50 contig length (Table 12) [118]. The assembled contigs of each sample were aligned against HMP urogenital tract reference genomes using ABACAS with default parameters, except 80% minimum percent identity [119]. The contigs that shared at least 80% identity with the HMP

urogenital tract reference genomes were called “assembly #1” and were combined with the following “assembly #2” and “assembly #3” for functional gene annotation.

3.3.7.3 Assembly #2 and assembly #3

The reads that were not used in “assembly #1” were assembled into contigs as previously described [103]. Briefly, the paired-end reads were digital normalized to k-mer coverage (defined as the number of times a k-mer was found in a read) of 20 using normalize-by-median.py in khmer v1.3. [120]. Highly repetitive k-mers, which could belong to multiple species, were removed using filter-below-abund.py with C=100 in khmer v1.3. [120]. The resulting paired-end reads were *de novo* assembled using Ray Meta at k=23 [121]. Scaffold coverage and length were calculated using Bowtie 2 [117] and samtools v0.1.18 [122]. Tetranucleotide frequency, GC content, and conserved marker genes were identified as described [103]. The binning of scaffolds into genomes was performed as described [103]. Scaffolds whose sequencing depths were bigger than 20X in any two subjects and which were classified by essential single copy gene were chosen (Table 13). The unassembled reads associated with these scaffolds were extracted and reassembled using SOAPdenovo v2.04 at k=23, thus, resulting new assembly [118]. The new assemblies were called “assembly #2”. The unassembled reads that were not used in “assembly #1” or “assembly #2” were *de novo* assembled using Ray Meta at k=23 [121]. This assembly was called “assembly #3”. “Assembly #2” and “assembly #3” were combined with “assembly #1” for the taxonomic and functional annotation of genes.

3.3.7.4 Functional annotation of genes

Assemblies 1, 2, and 3 were combined for each sample for functional annotation of genes. A total of 35 Mbp of contigs (>= 300 bp) were generated in the combined contigs from five samples with a longest contig length of 255 Kbp (Table 14). Microbial genes were also searched for in all contig sequences longer than 300 bp using the Prokaryotic Dynamic Programming Genefinding Algorithm (Prodigal) [123]. The predicted genes of all samples were combined to create a gene set. To build a non-redundant gene set, genes with 95% sequence identity and aligned length covered over 90% of the sequence were clustered together using cd-hit v4.6.1 [124]. To create a gene abundance matrix in which both pairs that mapped to the same gene would only be counted once, clean short reads in each sample were mapped back to the non-redundant gene set using bwa v0.7.5a [125] and samtools v0.1.18 [122]. Non-redundant genes were queried using DIAMOND [105] (default parameters except E=1e-5) against the KEGG database (v56) [106]. Only the best hits with at least 80% identity over 100 bp-aligned lengths were selected.

3.3.8. Network analysis

Vaginal discharge is the only BV symptom that was commonly found in all symptomatic BV patients. Yeoman *et al.* [1] found 33 metabolites correlated with vaginal discharge (Table 15). To explore the correlation among bacterial genus, KEGG orthologous groups, and the above 33 metabolites, Spearman's rank correlation coefficients and false discovery rate (FDR)-adjusted p-values associated with these correlation coefficients were generated using the "corr.test()" function in R. FDR-adjusted p-values < 0.05 were considered significant unless specifically stated. Significant correlations among bacterial

genus, KEGG orthologous groups, and discharge-associated metabolites were transformed into a visible interaction network in Cytoscape [130].

3.4 Results

3.4.1 Association between metabolites and KEGG orthologous groups abundance

Since vaginal discharge is the only common BV symptom found in symptomatic BV patients in our study, we only studied metabolites correlated with vaginal discharge. Thirty-three metabolites were filtered from Yeoman *et al.* [1] (Table 15). To determine whether we can find functional genes associated with vaginal discharge-associated metabolites, Spearman's rank correlation coefficients were used to investigate associations between all KEGG orthologous groups and these metabolites. Thirty-one and thirty-four KEGG orthologous groups were significantly positively and negatively linked to discharge-associated metabolites, respectively (Figure 7 and Table 16). 2-Methyl-2-hydroxybutanoic acid that was the only metabolite revealed by Yeoman *et al.* [1] was positively correlated with discharge. In our study, we detected that 'malate dehydrogenase' (K00024), '5-amino-6-(5-phosphoribosylamino) uracil reductase' (K11752), and 'undecaprenyl-diphosphatase' (K06153) were positively linked to this metabolite (Figure 8 and Table 16; r=1, p=0). Yeoman *et al.* [1] revealed that 2-Ethyl-4-methyl-1,3-Dioxolane, which was reported to exhibit pleasant odor [131], was negatively associated with discharge. In our study, three KEGG orthologous groups were positively linked to this metabolite (Figure 8 and Table 16; r=1, p=0). Three KEGG orthologous groups were positively linked to lactic acid (Figure 8 and Table 16; r=1, p=0) which was negatively associated with discharge in Yeoman *et al.* [1].

3.4.2 Association between metabolites and bacterial genera abundance

To determine whether we can find bacteria associated with vaginal discharge-associated metabolites, Spearman's rank correlation coefficients were used to investigate associations between all bacterial genera and discharge-associated metabolites. Fourteen and two bacterial genera were significantly positively and negatively linked to discharge-associated metabolites, respectively (Figure 7 and Table 16). Yeoman *et al.* [1] revealed that *Mobiluncus* was positively correlated with 2-Methyl-2-hydroxybutanoic acid. However, we did not find the correlation between *Mobiluncus* and 2-Methyl-2-hydroxybutanoic acid in our study. Instead, ten bacterial genera were positively associated with 2-Methyl-2-hydroxybutanoic acid in our study (Figure 9 and Table 16; r=1, p=0).

3.4.3 Association between KEGG orthologous groups and bacterial genera abundance

To better understand how vaginal bacteria generate discharge-associated metabolites, Spearman's rank correlation coefficients were used to investigate associations between previously found KEGG orthologous groups and bacterial genera. Twelve bacterial genera were significantly positively linked to eleven KEGG orthologous groups (Figure 7 and Table 16; r=1, p=0). Among these twelve bacterial genera and eleven KEGG orthologous groups, ten bacterial genera were positively associated with the metabolites ‘malate dehydrogenase’ (K00024), ‘5-amino-6-(5-phosphoribosylamino) uracil reductase’ (K11752), and ‘undecaprenyl-diphosphatase’ (K06153) (Figure 10 and Table

16; r=1, p=0). These three KEGG orthologous groups and ten bacterial genera were directly and positively linked to 2-Methyl-2-hydroxybutanoic acid, which was the only metabolite revealed by Yeoman *et al.* [1] that positively correlated with discharge (Figure 8-9 and Table 16; r=1, p=0). These suggest that these ten bacterial genera can serve as potential biomarkers to improve the development of therapeutic targets.

3.4.4 Annotation of functional genes involved in the generation of discharge-associated metabolites

To identify the most likely bacterial species that possess the functional genes involved in the generation of discharge-associated metabolites, these gene sequences were queried using BLASTN [98] with E=1e-5 against nucleotide collection reference database on July 10, 2015. Only the best hits with 100% query coverage and at least 80% identity were regarded as reliable hits. Under the criteria, 56 out of 408 genes were assigned to specific bacterial strains (Table 17). None of them belongs to genus *Aquabacterium*, *Chryseobacterium*, *Elizabethkingia*, *Finegoldia*, *Morganella*, *Negativicoccus*, *Phenylobacterium*, *Propionibacterium*, *Serratia*, and *Siphonobacter* that are potential therapeutic targets. This finding may result from the fact that most genomes of bacterial strains are not sequenced and therefore were not present in the reference sequence database.

3.5 Discussion

Studying the connection among bacteria, functional genes, and metabolites can lead to a better understanding of how pathogens initiate BV. Yeoman *et al.* [1] and Srinivasan *et al.*

[53] connected vaginal bacteria with BV symptom-associated metabolites and found vaginal bacteria that may be responsible for BV symptoms. Our study further investigated how vaginal bacteria produced BV symptom-associated metabolites by integrating the analysis of predicted gene functions. Our study found that the bacteria *Aquabacterium*, *Chryseobacterium*, *Elizabethkingia*, *Finegoldia*, *Morganella*, *Negativicoccus*, *Phenylobacterium*, *Propionibacterium*, *Serratia*, and *Siphonobacter* were directly or indirectly correlated with the metabolite '2-Methyl-2-hydroxybutanoic acid' via the enzymes 'malate dehydrogenase' (K00024), '5-amino-6-(5-phosphoribosylamino) uracil reductase' (K11752), and 'undecaprenyl-diphosphatase' (K06153). This suggests that these ten bacteria may produce the metabolite '2-Methyl-2-hydroxybutanoic acid' through these three functional genes. Future studies to determine whether or how these ten bacteria produce the metabolite '2-Methyl-2-hydroxybutanoic acid' through these three functional genes are needed.

We used the subset of study population in Yeoman *et al.* [1] to study the correlation between vaginal bacteria and vaginal discharge-associated metabolites like Yeoman and his colleagues did in their study. We did not find the same vaginal bacterium that was correlated with 2-Methyl-2-hydroxybutanoic acid, the only metabolite positively correlated with vaginal discharge. However, we found that *Aquabacterium*, *Chryseobacterium*, *Elizabethkingia*, *Finegoldia*, *Morganella*, *Negativicoccus*, *Phenylobacterium*, *Propionibacterium*, *Serratia*, and *Siphonobacter* were positively associated with 2-Methyl-2-hydroxybutanoic acid. The differences observed between these two studies could be due to two factors: (1) Due to small sample size (n=5), the

Spearman's rank correlation coefficient was used to investigate the correlations between bacterial genera and metabolites in this study. However, Pearson's correlation coefficient was used in Yeoman *et al.* [1]; (2) in this study, sequences collected from cervix, fornix, and vaginal canal by swabs, scrapings, and lavage were combined to be assigned to bacterial genera. However, Yeoman *et al.* [1] only used sequences collected from the vaginal canal by lavage. Sampling position and method may greatly affect analysis results: Kim *et al.* [129] showed that the vaginal microbiome was heterogeneous and three different sampling methods (swabbing, scraping, lavaging) characterized three considerably different vaginal microbiome profiles.

One of the limitations of this study is the small sample size. The limitation prevents us from finding more functional genes that are significantly correlated with vaginal bacteria and vaginal discharge-associated metabolites. The limitation also prevents us from finding the correlation among bacteria, functional genes, and other BV symptom-associated metabolites. The second limitation of this study is that samples collected by swabs, scrapings, and lavage were used for annotation of bacterial genera. However, only lavage samples were used for annotation of functional genes and measurement of metabolites.

3.6 Conclusions

Studying the correlation among vaginal bacteria, functional genes, and BV symptom-associated metabolites has allowed us to better understand how pathogens initiate BV. The bacteria *Aquabacterium*, *Chryseobacterium*, *Elizabethkingia*, *Finegoldia*,

Morganella, *Negativicoccus*, *Phenylobacterium*, *Propionibacterium*, *Serratia*, and *Siphonobacter* may be involved in the generation of vaginal discharge by producing the metabolite 2-Methyl-2-hydroxybutanoic acid via the functional genes ‘malate dehydrogenase’ (K00024), ‘5-amino-6-(5-phosphoribosylamino) uracil reductase’ (K11752), ‘undecaprenyl-diphosphatase’ (K06153). This suggests that these ten bacteria are potential therapeutic targets.

CHAPTER 4. A METAGENOMIC APPROACH TO CHARACTERIZE THE VAGINAL MICROBIOME THROUGHOUT PREGNANCY AND IN THE POSTPARTUM PERIOD

4.1 Abstract

Healthy women's vaginas are mainly dominated by *Lactobacilli* at reproductive age. The predominance of *Lactobacilli* is coincident with the rise in estrogen levels and glycogen content in the vaginal epithelium. Estrogen levels increase gradually during pregnancy, reaching a maximum near term, and drop down quickly during labor. Glycogen is converted to lactic acid mainly by *Lactobacilli*. Therefore, we hypothesized that the abundance of *Lactobacilli* and functional genes for metabolic activity in *Lactobacilli* would increase as pregnancy progresses and decrease following parturition. Vaginal samples were collected from 5 Caucasian women with and without a history of preterm birth at 8–12, 17–21, 26–30, and 35–38 weeks of gestation, during labor, and at 6 weeks post-partum. All women lived in Rochester, MN, USA from 23 to 43 years. The samples were collected by swabs from the posterior fornix and cervix. We used a metagenomic approach to classify sequences to bacterial species, GroopM [102] and CheckM [104] to recover the *Lactobacillus* population genomes, and DIAMOND [105] to blast the predicted genes in the *Lactobacillus* population genomes against the KEGG database (v56) to classify sequences to metabolic pathways. The majority of vaginal bacteria (defined as more than half of the vaginal bacteria) were *Lactobacillus* species in 17 out of 20 samples collected during pregnancy from women who delivered at term and who did not have a previous preterm birth history. On the other hand, for women who had a previous preterm

birth history but still delivered at term, the majority of vaginal bacteria were not *Lactobacillus* species in 10 out of 12 samples collected during pregnancy. Among these ten samples, the majority of vaginal bacteria in seven samples were *Gardnerella vaginalis*. Seven out of nine vaginal microbiomes collected from women with and without a previous preterm birth history were not dominated by any *Lactobacillus* species at 6 weeks post-partum. The relative abundance of *Lactobacilli*, the number of short reads participating in the recovery of the *Lactobacillus* population genomes, and the number of short reads predicted to participate in the metabolic activity in the *Lactobacillus* population genomes were higher in the samples collected at 35–38 weeks of gestation than in the samples collected at post-partum in all women with term birth. The relative abundance of *Lactobacilli*, the number of short reads participating in the recovery of the *Lactobacillus* population genomes, and the number of short reads predicted to participate in the metabolic activity in the *Lactobacillus* population genomes did not increase for any woman as pregnancy progressed from 8 weeks of gestational age to 30 weeks of gestational age. We conclude that the vaginal microbiomes of women who delivered at term and who did not have a previous preterm birth history were generally dominated by *Lactobacillus* species during pregnancy. The vaginal microbiomes of women who delivered at term but had a previous preterm birth history were more diverse than those of women who delivered at term and who did not have a previous preterm birth history.

4.2 Introduction

During the course of long-term evolutionary history, vaginal microorganisms have formed a mutualism relationship with humans. In this relationship, the host provides nutrients, like glycogen, to vaginal microorganisms to support their growth. In turn, vaginal microorganisms inhibit the colonization of pathogens. Perturbations in the vaginal microbiome have been associated with serious sequelae, such as bacterial vaginosis (BV), aerobic vaginitis, preterm birth, and miscarriage [3,7,38,51,132,133]. BV is one of the most common, major problems impacting women's reproductive health worldwide [22]. The prevalence of BV in the United States is about 29.2% or approximately 21.2 million women of reproductive age [23] and is associated with serious sequelae, such as pregnancy complications [3,27,37,38]. The United States annually spends nearly $1 billion on treating BV-related pregnancy complications [39]. Preterm birth is another major problems impacting women's reproductive health worldwide [73]. Every year, about 15 million babies are born preterm in the world [74]. The societal cost for preterm birth in the United States of America alone is more than $26.2 billion in 2005 [38].

Healthy women's vaginas are mainly dominated by *Lactobacilli* at reproductive age [134]. For example, Ravel *et al.* [9] used 16S-rRNA based approaches to study the vaginal microbiomes of 396 non-pregnant healthy American women representing four ethnic groups (white, black, Hispanic, and Asian). They showed that 73% of the non-pregnant healthy women's vaginal microbiomes were dominated by one species of *Lactobacillus* [9]. The same research group found that 94% of the healthy pregnant women's vaginal microbiomes are also dominated by one species of *Lactobacillus* [10]. The potential

mechanisms supporting the domination of *Lactobacilli* in vagina are: (1) they can generate lactic acid to maintain vagina acidity and inhibit the growth of pathogens [6,7,134,135]; (2) they can produce hydrogen peroxide and bacteriocins to kill pathogens in the vagina [6,7,134,135]; (3) they can competitively adhere to the vaginal epithelial cell receptors and form barrier to inhibit the adherence of pathogens [134,135]; and (4) they can stimulate host defense mechanisms to protect the host from the colonization of pathogens [134,135].

Researchers found that the predominance of *Lactobacilli* in the vaginal microbiomes of women of reproductive age is related to higher level of estrogen during reproductive years compared with other times in a woman's life [6,11]. Hickey *et al.* [6] and Farage *et al.* [11] showed that during puberty and reproductive years, estrogen production rises, resulting in an increase in the thickness of the vaginal epithelium, an increase of glycogen content in the vaginal epithelium, an acidic vaginal pH value, and the predominance of *Lactobacilli* in the vagina. In early childhood, the estrogen levels are low [6,11]. Low estrogen levels lead to the thinning of the vaginal epithelium, a decrease of glycogen content in the vaginal epithelium, a nearly neutral vaginal pH value, and the predominance of other various anaerobic bacteria in the vagina [6,11]. At postmenopause, estrogen levels decrease, causing a decrease of glycogen content in the vaginal epithelium and a nearly neutral vaginal pH value [6,11]. The vaginal microbiome shifts from a *Lactobacilli* dominated niche to poly-microbial communities during this time [6,11].

Estrogen levels are also reported to change the vaginal microbiome composition throughout the menstrual cycle and during pregnancy [3,12]. At menses, estrogen levels are

at their lowest in the menstrual cycle and non-*Lactobacillus* species increase compared with other times in the menstrual cycle [12]. The vaginal microbiomes of pregnant women were reported to be significantly different from those of non-pregnant women [10]. The vaginal microbiomes during pregnancy were also reported to significantly differ from those at 6 weeks post-partum [13]. Vaginal microbiomes during pregnancy have more *Lactobacillus* species compared with when not pregnant [10,13-15]. Estrogen level increases gradually during pregnancy, reaching a maximum near term, and drop down quickly during labor [107]. For example, the estrogen levels during pregnancy are 100-1000 fold higher than those at 1-week postpartum [13,16,17]. High estrogen levels increase glycogen content in the vaginal epithelium and promote the growth of *Lactobacillus* species [11,18].

The connection between the vaginal glycogen content, the colonization of *Lactobacillus* in vagina, and the vaginal pH value has also been reported by Mirmonsef *et al.* [136]. Mirmonsef *et al.* [136] revealed that the vaginal pH value is lower and the *Lactobacillus* colonization is higher in vaginal samples with high vaginal glycogen levels than vaginal samples with low vaginal glycogen levels. The potential mechanisms supporting the coincidence among the glycogen content, *Lactobacillus* colonization, and vaginal pH value are (1) glycogen is anaerobically converted to lactic acid; (2) *Lactobacilli* are the main component of lactic acid bacteria group; and (3) the function of lactic acid is to lower pH.

Therefore, we hypothesized that the abundance of *Lactobacilli* and functional genes for metabolic activity in *Lactobacilli* would increase as pregnancy progresses and decrease

following parturition. The objectives of this study are to determine whether the abundance of *Lactobacilli* and functional genes for metabolic activity in *Lactobacilli* increase as pregnancy progresses and decrease following parturition. We followed a metagenomic approach using vaginal samples from women with and without a history of preterm birth at 8–12, 17–21, 26–30, and 35–38 weeks of gestation, during labor, and at 6 weeks post-partum to determine the taxonomic and functional composition of the vaginal microbiome collected at different time points.

4.3 Methods

4.3.1. Ethics statement

This study was approved and reviewed by the Mayo Clinic Institutional Review Board, USA (IRB #10-006257). Informed written consent was obtained from all study participants prior to sample collection.

4.3.2. Sample collection

Vaginal samples were collected from women with (IDs: 110, 117, 119, 120, 125) and without (IDs: 201, 202, 203, 205, 206) a history of preterm birth at 8–12, 17–21, 26–30, and 35–38 weeks of gestation, during labor, and at 6 weeks post-partum. The inclusion and exclusion criteria for participation in the study were the same as described in a previous study [14]. The sample descriptions are given in Table 18. Subjects 201 and 205 gave birth at 34 4/7 and 36 5/7 weeks of gestation, respectively. The remaining eight subjects gave birth after 37 weeks of gestation. The samples were collected by swabs that were collected by the obstetrician or certified nurse midwife from the posterior fornix and

cervix and placed in a NAT (Nucleic Acid Transport, CentraCare Laboratory Services, St. Cloud, MN) collection tube. After collection, the samples were stored at -80°C until use.

4.3.3. Sample processing

The vaginal swabs were vigorously agitated to dislodge the bacterial cells. The cells were pelleted by centrifugation at 10,000 g for 10 min. Bacterial genomic DNA was extracted using a MoBio PowerSoil DNA Isolation Kit (MoBio Laboratories Inc., Carlsbad, CA, USA) following the manufacturer's instructions. Concentration of DNA was determined using High Sensitivity Qubit (Life Technologies Corporation, Carlsbad, CA, USA). The integrity and fragment size of the extracted DNA were determined by 1% agarose gel electrophoresis. All DNA samples were sequenced on Illumina HiSeq 2000 at the W. M. Keck Center for Biotechnology at the University of Illinois at Urbana-Champaign.

4.3.4. Metagenomic sequencing data analysis and statistics

4.3.4.1 Reads trimming

The qualities of the raw Illumina sequencing data were first assessed using FastQC v0.10.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The first few bases of each sequence, which showed a bias in the FastQC report, were trimmed using the FASTX toolkit [115]. Leading and trailing low quality bases (Q score below 20) and sequences shorter than 50 nucleotides and with an average quality score < 20 were removed using trimmomatic v0.30 [116]. This quality filtering resulted in the quality values of all bases being higher than 28 in the FastQC report. To remove potential

contamination with human DNA, the remaining reads were further mapped against human genome reference sequence (NCBI build 37 p13) with one mismatch permitted per read in Bowtie 2 [117]. A total of 3,406,720,858 raw short reads from fifty-three samples were generated with an average number of reads of 64,277,752 per sample (Table 19). After quality filtering and human DNA contamination removing, a total of 469,315,112 short reads remained with an average number of reads of 8,855,002 per sample (Table 19).

4.3.4.2 Taxonomic analyses

We used two different approaches to assign the resulting clean reads to bacteria: (1) direct assignment of unassembled reads and (2) indirect assignment of unassembled reads through contigs. For direct assignment, bacterial species were identified using Taxonomic identification and phylogenetic profiling (TIPP) [100]. TIPP uses 40 phylogenetic marker genes that are single copy and universally present in all bacteria to identify bacterial species and determine their abundance from metagenomic shotgun sequencing data [100]. For indirect assignment, metagenomic short reads were first assembled into contigs. Bacterial species were then identified from contigs.

4.3.4.2.1 Assembly #1

To obtain a reliable assembly, the resulting clean reads from samples collected at different time point were combined by subject. The combined metagenomic data were aligned against the HMP urogenital tract reference genomes downloaded on January 10, 2015 (Table 3). The read alignments with mapping qualities > 20 were considered to be

true alignments. The reads whose breadth of coverage (defined as the percentage of bases covered in the entire reference genome) was more than 2% at least one in sample were further assembled to contigs using SOAPdenovo v2.04 with k-mer 31 (Table 20) [118]. The assembled contig for each bacterium were combined by subject. The combined assembly were called “assembly #1” and were combined with the following “assembly #2” for the further analyses.

4.3.4.2.2 Assembly #2

The reads that were not used in “assembly #1” were *de novo* assembled using Ray Meta at k=23 [121]. This assembly was called “assembly #2”. “Assembly #2” was combined with “assembly #1” for the further analyses

4.3.4.2.3 Taxonomic classification of contigs

A total of 76 Mbp of contigs (>= 500 bp) were generated in the combined contigs from fifty-three samples with longest contig length of 536 Kbp (Table 21). Microbial genes were also searched for in all contig sequences longer than 500 bp using the Prokaryotic Dynamic Programming Genefinding Algorithm (Prodigal) [123]. The predicted genes of all subjects were combined to create a gene set. To build a non-redundant gene set, genes with 95% sequence identity and aligned length covered over 90% of the sequence were clustered together using cd-hit v4.6.1 [124]. To create a gene abundance matrix in which both pairs that mapped to the same gene would only be counted once, clean short reads in each sample were mapped back to the non-redundant gene set using bwa v0.7.5a [125] and samtools v0.1.18 [122]. The non-redundant genes were queried using BLASTN [98] (default

parameters except E=1e-5) against the HMP urogenital tract reference database downloaded on January 10, 2015 (Table 3) and the NCBI complete bacteria genome, downloaded on September 6, 2014. Only the best hits with at least 80% identity over 100 bp-aligned lengths were selected to build the species abundance matrix.

4.3.4.3 Population genomes analyses

To obtain coverage profiles, the clean short reads from each sample were separately mapped to the corresponding "assembly #1 and 2" using Bowtie 2 [117]. The resulting SAM files were converted to sorted-indexed BAM files using samtools v0.1.18 [122]. The coverage profiles (sorted-indexed BAM files) and combined contigs (>= 500 bp) were parsed into GroopM v0.3.4 [102] to bin contigs, resulting in the generation of a total of 111 population genomes bins across ten subjects (Table 21). After estimating the completeness and contamination of each population genomes bin using CheckM [104], twelve high-quality population genomes were recovered with at least >88% completeness and <9% contamination (Table 22 and Figure 11). Interestingly, all of the twelve high-quality population genomes were classified as *Lactobacillus* using CheckM [104] (Table 22). Moreover, at least one high-quality *Lactobacillus* population genomes was recovered from each person (Table 22). To determine whether the high-quality *Lactobacillus* population genomes was mainly contributed by the samples collected during pregnancy, the clean short reads in each sample were mapped against the corresponding *Lactobacillus* population genomes in Bowtie 2 [117]. The absolute number of short reads in each sample mapped against the *Lactobacillus* population genomes was then divided by

the total number of clean reads and multiplied by 1,000,000 for the purposes of normalization.

To determine whether the functional genes for metabolic activity in *Lactobacilli* increases as pregnancy progresses and decrease following parturition, contigs participated in the *Lactobacillus* population genomes in each sample were predicted to microbial genes using Prokaryotic Dynamic Programming Genefinding Algorithm (Prodigal) [123]. The predicted genes were queried using DIAMOND [105] (default parameters except E=1e-5) against the KEGG database (v56) [106]. Only the best hits with at least 80% identity over 100 bp-aligned lengths were selected. The clean short reads in each sample were mapped against the predicted metabolic genes in Bowtie 2 [117]. The absolute number of short reads in each sample mapped against the predicted metabolic genes was then divided by the total number of clean reads and multiplied by 100,000,000 for the purposes of normalization.

4.4 Results

4.4.1 Taxonomic analysis

To determine whether *Lactobacillus* species are the majority vaginal bacteria during pregnancy but not at post-partum, unassembled reads were directly assigned to bacteria using TIPP [100] or indirectly assigned to bacteria via contigs. The two different methods generated similar results (Figure 12 and 13, Table 23 and 24). The majority of vaginal bacteria are defined as more than half of the bacteria in vagina. The majority of vaginal bacteria were *Lactobacillus* species in 17 out of 20 samples collected during pregnancy

from women who delivered at term and who did not have a previous preterm birth history (Figure 12 and 13, Table 23 and 24). Four out of five vaginal microbiomes collected from women who delivered at term and who did not have a previous preterm birth history were not dominated by any *Lactobacillus* species at 6 weeks post-partum (Figure 12 and 13, Table 23 and 24). Although the majority vaginal bacteria in subject 120 (without a previous preterm birth history) at the early state of pregnancy (8–12, 17–21, and 26–30 weeks of gestation for directly assignment method; 8–12 and 26–30 weeks of gestation for indirectly assignment method) were *Gardnerella vaginalis*, *Lactobacillus crispatus* became the majority vaginal bacterium at 35–38 weeks of gestation (Figure 12 and 13, Table 23 and 24). Therefore, the majority of vaginal bacteria in women who delivered at term and who did not have a previous preterm birth history were *Lactobacillus* species at 35–38 weeks of gestation (Figure 12 and 13, Table 23 and 24).

The above trends were not seen in the women with a previous preterm birth history (Figure 12 and 13, Table 23 and 24). The vaginal microbiomes of women with a previous preterm birth history were more diverse than those of women without a previous preterm birth history (Figure 12 and 13, Table 23 and 24). For women who had preterm delivery and a previous preterm birth history, the majority vaginal bacterium during pregnancy was *Lactobacillus iners* and *crispatus*, respectively (Figure 12 and 13, Table 23 and 24). However, for women who had a previous preterm birth history but still delivered at term, the majority of vaginal bacteria were not *Lactobacillus* species in 10 out of 12 samples collected during pregnancy (Figure 12 and 13, Table 23 and 24). Among these ten samples, the majority of vaginal bacteria in seven samples were *Gardnerella vaginalis*

(Figure 12, Table 23 and 24). The vaginal bacteria in women who had a previous preterm birth history but still delivered at term at 6 weeks post-partum were evenly distributed -- that is to say, the relative abundance of all vaginal bacteria was below 50% (Figure 12 and 13, Table 23 and 24). The vaginal microbiome of subject 205, who had preterm delivery and a previous preterm birth history, was dominated by *Lactobacillus crispatus* at 6 weeks post-partum (Figure 12 and 13, Table 23 and 24). Interestingly, the relative abundance of *Lactobacilli* at 35–38 weeks of gestation in women who had a previous preterm birth history but still delivered at term was below 50% (Figure 12 and 13, Table 23 and 24).

The vaginal microbiomes of five out of ten women persisted to be dominated by one species of *Lactobacillus* throughout pregnancy The remaining five women showed a transition from one majority vaginal bacterium to another majority vaginal bacterium throughout pregnancy (Figure 12 and 13, Table 23 and 24). The vaginal microbiomes of any women who had a previous preterm birth history but still delivered at term did not persist in any one bacterial species throughout pregnancy (Figure 12 and 13, Table 23 and 24).

To determine whether the relative abundance of *Lactobacilli* increases as pregnancy progresses and decrease following parturition, the relative abundances of all *Lactobacillus* species at 8–12, 17–21, 26–30, and 35–38 weeks of gestation and 6 weeks post-partum were added up. The relative abundances of *Lactobacilli* at 8–12, 17–21, 26–30, and 35–38 weeks of gestation were compared with those at 6 weeks post-partum. For

women without a history of preterm birth, the relative abundances of *Lactobacilli* were higher in the samples collected during pregnancy than in the samples collected at 6 weeks post-partum (Figure 14A and B, Table 23 and 24). However, for women with a history of preterm birth, this trend was only seen in subject 202 (Figure 14C and D, Table 23 and 24). The relative abundances of *Lactobacilli* were lower in the samples collected at 6 weeks post-partum than in the samples collected at 8–12 and 17–21 weeks of gestation in subjects 203 and 206 (Figure 14C and D, Table 23 and 24). The relative abundance of *Lactobacilli* in subject 205 did not have significantly change throughout pregnancy and post-partum (Figure 14C and D, Table 23 and 24). However, the relative abundance of *Lactobacilli* was higher in the samples collected at 26–30 and 35–38 weeks of gestation than in the samples collected at post-partum in all women who delivered at term (Figure 14, Table 23 and 24). Moreover, the relative abundance of *Lactobacilli* did not significantly increase in any of the ten women as pregnancy progressed from 8 weeks of gestational age to 30 weeks of gestational age (Figure 14, Table 23 and 24). These observations held true whether the relative abundances of *Lactobacilli* were determined from the metagenomic short reads or from the assembly.

4.4.2 Population genomes analysis

The above focuses on the genes of *Lactobacilli*. What happened to the *Lactobacillus* population genomes as a whole? To recover population genomes from metagenomes, we used GroopM to bin contigs into bins, where each contig belonged to a single genome or some closely related genomes [102]. Binning was performed based on the co-assembly of all samples in each subject. After estimating the completeness and contamination of each bin

using CheckM [104], the high-quality population genomes were members of the genus *Lactobacillus* (Table 22). To determine whether the high-quality *Lactobacillus* population genomes were mainly contributed by the samples collected during pregnancy, we counted and normalized the number of clean short reads in each sample that were used to assemble corresponding *Lactobacillus* population genomes. For women who delivered at term and who did not have a previous preterm birth history, the number of short reads participating in the recovery of the *Lactobacillus* population genomes was higher in the samples collected during pregnancy than in the samples collected at 6 weeks post-partum except subject 120 (Figure 15A and Table 19). However, for women with a history of preterm birth, this trend was only seen in subject 202 (Figure 15B and Table 19). The number of short reads participating in the recovery of the *Lactobacillus* population genomes from subjects 203 and 206 at 6 weeks post-partum were similar to those at 8–12 and 17–21 weeks of gestation (Figure 15B and Table 19). The number of short reads participating in the recovery of the *Lactobacillus* population genomes was higher in the samples collected at 6 weeks post-partum than in the samples collected during pregnancy in subject 205 (Figure 15B and Table 19). However, the number of short reads participating in the recovery of the *Lactobacillus* population genomes was higher in the samples collected at 35–38 weeks of gestation than in the samples collected at post-partum in all women who delivered at term (Figure 15 and Table 19). It did not increase in any of the ten women as pregnancy progressed from 8 weeks of gestational age to 30 weeks of gestational age (Figure 15 and Table 19).

To determine whether the number of short reads predicted to participate in the metabolic activity in the *Lactobacillus* population genomes increase as pregnancy progresses and decrease following parturition, open reading frames (ORFs) were predicted in contigs participating in the *Lactobacillus* population genomes in each sample, and then aligned to the genes in the KEGG pathways [106]. The number of short reads predicted to participate in the metabolic activity in the *Lactobacillus* population genomes were counted and normalized in each sample. The number of short reads predicted to participate in the metabolic activity in the *Lactobacillus* population genomes had the same pattern as the number of short reads participating in the recovery of the *Lactobacillus* population genomes (Figure 16 and Table 19). The number of short reads predicted to participate in the metabolic activity in the *Lactobacillus* population genomes was higher in the samples collected at 35–38 weeks of gestation than in the samples collected at post-partum in all women who delivered at term (Figure 16 and Table 19). It did not significantly increase in any of the ten women as pregnancy progressed from 8 weeks of gestational age to 30 weeks of gestational age (Figure 16 and Table 19).

4.5 Discussion

Our longitudinal study revealed that *Lactobacillus iners* or *crispatus* were enriched in the vaginal microbiomes collected during pregnancy from women who delivered at term and who did not have a previous preterm birth history. This is consistent with the results from a previous study from twelve women with a healthy pregnancy progression and outcome, which revealed that *Lactobacillus iners* or *crispatus* dominated in the vaginal microbiomes of these women [14]. In our study, whole genome shotgun sequencing was

used to characterize the vaginal microbiome profile in Caucasian women. In Walther-Antonio *et al.* [14], they used 16S rRNA gene sequencing to characterize the vaginal microbiome profile in Caucasian women. There was only an overlap of one subject between the two studies [14]. The results of both studies suggest that the vaginal microbiomes of Caucasian women who delivered at term and who did not have a previous preterm birth history are characterized by the enrichment of *Lactobacillus iners* or *crispatus* during pregnancy. The predominance of *Lactobacillus* spp. in the vaginal microbiomes of pregnant women is also observed in another two longitudinal studies [10,13] and one cross–sectional study [15].

Our longitudinal study revealed that the vaginal microbiomes of women who delivered at term and who did not have a previous preterm birth history became less *Lactobacillus* spp. enriched at post-partum compared with pregnancy. The finding agrees with a previous study revealing that vaginal microbiomes shift from more *Lactobacillus* species during pregnancy to less *Lactobacillus* species when not pregnant [13]. The shift is coincident with the rise in estrogen levels and glycogen content in the vaginal epithelium during pregnancy [11,13,16,17,107]. Glycogen in the vaginal epithelium provides a carbon source to support the growth of *Lactobacillus* spp. and keeps the vagina acidic to inhibit the colonization of other various bacteria through the generation of lactic acid [17].

The non-*Lactobacillus* dominated vaginal community has long been regarded as an unhealthy vaginal community that increases the risk of preterm birth. However, in our study, three out of ten women whose vaginal microbiomes collected during pregnancy

were not dominated by *Lactobacillus* delivered mature babies. These results suggest that the loss of *Lactobacillus* is not a reliable sign of preterm birth. In accordance with our finding, Hyman *et al.* [137] concluded that the relative absence of *Lactobacillus* and the presence of other various bacteria are not related to preterm birth [137]. Similar finding was also found by Romero *et al.* [138]. They showed that 27 out of 72 women who delivered at term had at least one vaginal microbiome collected during pregnancy that was not dominated by *Lactobacillus.* In our study, when compared with women who delivered at term and who did not have a previous preterm birth history, women who delivered at term but had a previous preterm birth history were more likely to have vaginal microbiomes not dominated by *Lactobacillus* throughout pregnancy. This implies that a previous history of preterm birth affects the distribution of vaginal bacteria.

Our study observed that vaginal microbiomes collected during pregnancy from women who had preterm delivery were dominated by one species of *Lactobacillus.* Similar finding was found by Hyman *et al.* [137]. They observed that vaginal microbiomes collected during pregnancy from two Caucasian women who delivered preterm were dominated by *Lactobacillus crispatus*. This suggests that the predominance of *Lactobacillus* does not guarantee term birth. However, previous studies demonstrated that the predominance of *Lactobacilli* in vagina is the signature of term pregnancy [14,15]. Therefore, it is possible that vaginal microbiome composition does not make a difference for preterm birth. For example, Romero *et al.* [138] revealed that the vaginal microbiomes of pregnant women who delivered preterm did not differ from those of pregnant women who delivered at term. It is also possible that the predominance of *Lactobacilli* is an important factor

indicating term birth. However, other things could go wrong, such as preterm premature rupture of membranes (PPROM), which we hypothesize was the cause of the preterm birth for the two women whose vagina dominated by one species of *Lactobacillus*. One of these two women (subject 201) presented to hospital with preterm labor at 34 4/7 weeks of gestation and her previous preterm delivery was due to PPROM at 33 weeks of gestation; the other (subject 205) presented to hospital at 36 5/7 weeks of gestation with premature rupture of membranes and she had a previous history of PPROM. PPROM is a rupture of fetal membrane before 37 weeks of gestation and accounts for 25% of preterm birth [75]. The cause of PPROM is still unknown. However, infection/inflammation, decidual bleeding (abruption), uterine overdistention, genetic predispositions, and cigarette smoking have been reported to be associated with PPROM [139].

Our study observed the transition from one majority vaginal bacterium to another majority vaginal bacterium throughout pregnancy in five out of ten women. This finding is consistent with two studies of Romero and his colleagues [10,138]. In these two studies, they observed the shift of vaginal community state types (CSTs) throughout pregnancy in both women who delivered at term and those that delivered preterm [10,138]. The shift could be from CST dominated by one species of *Lactobacillus* to CST dominated by another species of *Lactobacillus* or from CST dominated by one species of *Lactobacillus* to CST not dominated by *Lactobacillus* or from CST not dominated by *Lactobacillus* to CST dominated by one species of *Lactobacillus* [10,138].

Our study revealed that for subjects 203 and 206 who had a previous preterm birth history but still delivered at term, the relative abundance of *Lactobacilli*, the number of short reads participating in the recovery of the *Lactobacillus* population genomes, and the number of short reads for predicted metabolic genes in the *Lactobacillus* population genomes were similar or lower at the early stage of pregnancy compared to post-partum. However, those values increased at the later stage of pregnancy. This may result from the increase of levels of circulating glucose in pregnant women in late pregnancy [140]. In late pregnancy, increase of levels of circulating glucose may provide more substrate to *Lactobacillus* species to be fermented to lactate.

Our study found that the abundance of *Lactobacilli* and functional genes for metabolic activity in *Lactobacilli* did not significantly increase in any of the ten women as pregnancy progressed from 8 weeks of gestational age to 30 weeks of gestational age. For six out of ten women whose vaginal microbiomes were dominated by *Lactobacilli* during pregnancy, the increase of the abundance of *Lactobacilli* and functional genes for metabolic activity in *Lactobacilli* may have happened before 8 weeks of gestational age. For women of reproductive age, the level of estrogen is at its highest during the middle of the menstrual cycle when women are ovulating [12]. The presence of a fertilized egg will trigger the increased production of estrogen which makes the lining of the uterus grow and thicken to prepare for the embryo implantation. High estrogen levels increase glycogen content in the vaginal epithelium and promote the growth of *Lactobacillus* species [11,18]. For the remaining four women whose vaginal microbiomes were not

dominated by *Lactobacilli* during pregnancy, other factors except estrogen probably influenced the vaginal microbiome composition.

One of the limitations of this study is that we did not collect vaginal samples before 8 weeks of gestation during pregnancy, before pregnancy, or during the middle of the menstrual cycle. Therefore, we do not know how the abundance of *Lactobacilli* changes in very early pregnancy, the vaginal microbiomes composition when these ten women were not pregnant or breast-feeding, and whether the vaginal microbiome composition during the middle of the menstrual cycle was similar to that in the early pregnancy. The second limitation of this study is that we did not measure estrogen level, vaginal glycogen content, or vaginal pH value. Therefore, it is difficult to establish the direct relationship among estrogen level, vaginal glycogen content, colonization of *Lactobacillus* in vagina, and vaginal pH value. The third limitation of this study is the small sample size. The limitation prevents us from drawing statistical conclusions.

4.6 Conclusions

The vaginal microbiomes of women who delivered at term and who did not have a previous preterm birth history were generally dominated by *Lactobacillus* species during pregnancy. The vaginal microbiomes of women who delivered at term but had a previous preterm birth history were more diverse than those of women who delivered at term and who did not have a previous preterm birth history. The abundance of *Lactobacilli* and functional genes for metabolic activity in *Lactobacilli* did not increase as pregnancy progressed from 8 weeks of gestational age to 30 weeks of gestational age. The

abundance of *Lactobacilli* and functional genes for metabolic activity in *Lactobacilli* were higher in the samples collected in late pregnancy than in the samples collected at post-partum in all women with term birth.

CHAPTER 5. DISSERTATION SUMMARY AND FUTURE RESEARCH

5.1 Dissertation summary

BV is one of the major problems impacting women's reproductive health worldwide. It is caused by the perturbations in the vaginal microbiome. The etiology of BV is multifactorial. Because of this, the diagnosis and treatment of BV are effective. To better understand the mechanism of BV, we have to understand what bacteria cause BV and how they do it. Most existing research has focused on what bacteria cause BV -- a minority of existing research focuses on what the pathogens are doing. A metagenomic approach can answer both the question of what bacteria are in the microbiome and provide insight into what they are doing. Therefore, in the first and second projects (Chinese and US BV project) in this dissertation, I used a metagenomic approach to understand the relationship between the vaginal microbiome and bacterial vaginosis (BV).

The hypothesis of Chinese BV project is that sialidase can serve as a biomarker for BV (chapter 2). The rationales behind the hypothesis are (1) It has been demonstrated that sialidase secreted by vaginal bacteria degrades mucin by hydrolyzing sialic acid from mucin [66,67]. The function of mucin is to protect the vaginal epithelium from the colonization of pathogens [66,67]. Sialidases therefore assist the adherence, colonization, and invasion of pathogens [67]. (2) Studies have shown that the level of sialidase activity is significantly higher in BV patients compared with that in women without BV [61,62,68-70]. (3) *Prevotella*, *Gardnerella*, and *Bacteroides* species, which are reported to be associated

with BV, were reported to have the ability to secrete sialidase [59,61-63,65]. In my first study, there were three objectives to address. First, I determined whether sialidase activity is elevated in Chinese BV patients. Second, I determined whether the relative abundance value of sialidase-encoding genes increases in Chinese BV patients. Finally, I determined whether the relative abundances of *Prevotella*, *Gardnerella,* and *Bacteroides* species are higher in Chinese BV patients.

If sialidase can serve as a biomarker for BV, sialidase activity should be elevated in Chinese BV patients but not in Chinese women without BV. To determine this, we used the BVBlue sialidase test to measure the sialidase activity in each Chinese sample. Elevated sialidase activity was detected in all BV patients but not in any women without BV. Therefore, the test was able to distinguish between women with and without BV. To further determine whether the sialidase test could separate women with and without BV, we examined the microbiome diversity and microbial community structure of women with and without BV. The vaginal microbiomes of BV patients were more diverse than those of women without BV. Bray-Curtis dissimilarity-based PCoA plot exhibited a disparate clustering between women with and without BV. Therefore, women with and without BV harbored distinctly different vaginal microbiomes.

The sialidase enzymes secreted by vaginal bacteria are encoded by functional genes in those bacteria. To determine whether the relative abundance value of sialidase-encoding genes increases in BV patients, we analyzed whole genome shotgun (WGS) sequences in microbial samples from Chinese women with and without BV and identified functional

genes that were more abundant in Chinese BV patients. The microbiomes of all Chinese women lacked gene sequences encoding a sialidase enzyme. This finding may be due to inadequate sequencing depth.

Since sialidase-encoding genes were not found in all samples, what other genes could differentiate the BV patients from non-BV patients? To find an alternative to sialidase-encoding genes, we compared all predicted enzymes of Chinese women with BV with those of Chinese women without BV. We found that thiol-activated cytolysin (K11031), which is a toxin, and putative peptidoglycan lipid II flippase (K03980), which is virulence factor, were more abundant in BV patients.

To determine whether *Prevotella*, *Gardnerella*, and *Bacteroides* species are more abundant in BV patients, we analyzed 16S-rRNA gene and WGS sequences in microbial samples from Chinese women with and without BV and identified bacterial genera and species that were more abundant in Chinese BV patients. Analysis of 16S rRNA gene and WGS sequence data revealed that *Prevotella* and *Gardnerella* genera and species were more abundant in BV patients.

The conclusions of Chinese BV project were that the sialidase test could separate diseased vaginal microbiomes from non-diseased ones. *Prevotella* and *Gardnerella* species were associated with high sialidase levels.

The hypothesis of US BV project is that BV symptom-associated bacteria can serve as potential biomarkers to improve the development of therapeutic targets (chapter 3). The rationales behind the hypothesis are (1) Yeoman *et al.* [1] and Srinivasan *et al.* [53] revealed that women with and without BV have distinct metabolite profiles and found metabolites that are correlated with BV symptoms: discharge, odor, itching, and pain. (2) The same studies found vaginal bacteria that were correlated with BV symptom-associated metabolites. However, neither study connected functional genes to vaginal bacteria and metabolites. BV symptom-associated metabolites are encoded by functional genes in vaginal bacteria. Therefore, studying the correlation among bacteria, functional genes, and metabolites could help better understand how pathogens initiate BV. To help test this hypothesis, we took a subset of the population from Yeoman *et al.* [1] and examined the correlation among vaginal bacteria, vaginal discharge-associated metabolites, and functional genes. We then pursued the following two objectives: (1) Identify bacteria and functional genes associated with metabolites that Yeoman *et al.* [1] found to be correlated with vaginal discharge, and (2) Identify which identified bacteria were correlated with the identified functional genes.

Vaginal discharge-associated metabolites are encoded by functional genes in vaginal bacteria. Moreover, Yeoman *et al.* [1] found vaginal bacteria that were correlated with discharge-associated metabolites. To determine whether we can find bacteria and functional genes associated with vaginal discharge-associated metabolites, we analyzed 16S-rRNA gene and WGS sequences in microbial samples from the US women with and without BV and performed a Spearman's rank correlation on bacterial genera, functional

genes, and discharge-associated metabolites. The enzymes ‘malate dehydrogenase’ (K00024), ‘5-amino-6-(5-phosphoribosylamino) uracil reductase’ (K11752), and ‘undecaprenyl-diphosphatase’ (K06153), and the bacteria *Aquabacterium*, *Chryseobacterium*, *Elizabethkingia*, *Finegoldia*, *Morganella*, *Negativicoccus*, *Phenylobacterium*, *Propionibacterium*, *Serratia*, and *Siphonobacter* were significantly positively linked to the metabolite ‘2-Methyl-2-hydroxybutanoic acid’.

To better understand how vaginal bacteria generate discharge-associated metabolites, we performed a Spearman’s rank correlation test between bacteria and functional genes. The bacteria *Aquabacterium*, *Chryseobacterium*, *Elizabethkingia*, *Finegoldia*, *Morganella*, *Negativicoccus*, *Phenylobacterium*, *Propionibacterium*, *Serratia*, and *Siphonobacter* were significantly positively linked to the enzymes ‘malate dehydrogenase’ (K00024), ‘5-amino-6-(5-phosphoribosylamino) uracil reductase’ (K11752), and ‘undecaprenyl-diphosphatase’ (K06153). Finally, to identify the most likely bacterial strains that possess the functional genes involved in the generation of discharge-associated metabolites, these gene sequences were queried using BLASTN [98] against the nucleotide collection reference database. A small percentage of genes were assigned to specific bacterial strains. Among the classified bacterial strains, none of them belongs to genus *Aquabacterium*, *Chryseobacterium*, *Elizabethkingia*, *Finegoldia*, *Morganella*, *Negativicoccus*, *Phenylobacterium*, *Propionibacterium*, *Serratia*, and *Siphonobacter*. This finding may result from the fact that most genomes of bacterial strains are not sequenced and therefore were not present in the reference sequence database.

We found that the bacteria *Aquabacterium*, *Chryseobacterium*, *Elizabethkingia*, *Finegoldia*, *Morganella*, *Negativicoccus*, *Phenylobacterium*, *Propionibacterium*, *Serratia*, and *Siphonobacter* may be involved in the generation of vaginal discharge by producing the metabolite '2-Methyl-2-hydroxybutanoic acid' via the enzymes 'malate dehydrogenase' (K00024), '5-amino-6-(5-phosphoribosylamino) uracil reductase' (K11752), and 'undecaprenyl-diphosphatase' (K06153). This suggests that these ten bacteria are potential therapeutic targets.

In my third project (pregnancy project), I used a metagenomic approach to characterize the vaginal microbiome throughout pregnancy and in the postpartum period. The hypothesis of this project is that the abundance of *Lactobacilli* and functional genes for metabolic activity in *Lactobacilli* will increase as pregnancy progresses and decrease following parturition. The rationales behind the hypothesis are (1) Vaginal microbiomes are mainly dominated by *Lactobacilli* during pregnancy [10,13-15]; (2) The predominance of *Lactobacilli* in the vagina has been reported to be related to high estrogen levels and glycogen content in the vaginal epithelium [6,11]; (3) Estrogen levels increase gradually during pregnancy, reaching a maximum near term, and drop down quickly during labor [107]; (4) Glycogen is anaerobically converted to lactic acid in the vagina; and (5) *Lactobacilli* are the main component of lactic acid bacteria group. I formulated two objectives to address the hypothesis. The two objectives are to (1) Determine whether the abundance of *Lactobacilli* increases as pregnancy progresses and decreases following parturition; and (2) Determine whether the functional genes for metabolic activity in *Lactobacilli* increases as pregnancy progresses and decrease following parturition.

To realize these two objectives, we followed a metagenomic approach using vaginal samples from women with and without a history of preterm birth at 8–12, 17–21, 26–30, and 35–38 weeks of gestation, during labor, and at 6 weeks post-partum to determine the taxonomic and functional composition of the vaginal microbiome collected at different time points.

The vaginal microbiome composition was reported to differ during pregnancy compared to when not pregnant [10,13,14]. Vaginal microbiomes during pregnancy were reported to have more *Lactobacillus* species [10,13-15]. To determine whether *Lactobacillus* species are the majority vaginal bacteria during pregnancy but not at post-partum, unassembled reads were directly assigned to bacteria using TIPP [100] or indirectly assigned to bacteria via contigs. The majority of vaginal bacteria are defined as more than half of the bacteria in vagina. The majority of vaginal bacteria were *Lactobacillus* species in 17 out of 20 samples collected during pregnancy from women who delivered at term and who did not have a previous preterm birth history. On the other hand, for women who had a previous preterm birth history but still delivered at term, the majority of vaginal bacteria were not *Lactobacillus* species in 10 out of 12 samples collected during pregnancy. Among these ten samples, the majority of vaginal bacteria in seven samples were *Gardnerella vaginalis*. Seven out of nine vaginal microbiomes collected from women with and without a previous preterm birth history were not dominated by any *Lactobacillus* species at 6 weeks post-partum. These observations held true whether the relative abundance of

Lactobacillus species was determined from the metagenomic short reads or from the assembly.

Estrogen levels increase gradually during pregnancy, reaching a maximum near term, and drop down quickly during labor [107]. To determine whether the relative abundance of *Lactobacilli* increases as pregnancy progresses and decrease following parturition, the relative abundances of all *Lactobacillus* species at 8–12, 17–21, 26–30, and 35–38 weeks of gestation and 6 weeks post-partum were added up. The relative abundances of *Lactobacilli* at 8–12, 17–21, 26–30, and 35–38 weeks of gestation were compared with those at 6 weeks post-partum. The relative abundances of *Lactobacilli* were higher in the samples collected at 26–30 and 35–38 weeks of gestation than in the samples collected at post-partum in all women with term birth. The relative abundance of *Lactobacilli* did not increase in any women as pregnancy progressed from 8 weeks of gestational age to 30 weeks of gestational age.

The above focuses on the genes of *Lactobacilli*. What happened to the *Lactobacillus* population genomes as a whole? To determine whether the number of short reads participating in the recovery of the *Lactobacillus* population genomes increases as pregnancy progresses and decreases following parturition, we used GroopM to bin contigs into bins, where each contig belonged to a single genome or several closely related genomes [102]. Binning was performed based on the co-assembly of all samples in each subject. Then, we counted and normalized the number of clean short reads in each sample that were used to assemble corresponding *Lactobacillus* population genomes. The

number of short reads participating in the recovery of the *Lactobacillus* population genomes was higher in the samples collected at 35–38 weeks of gestation than in the samples collected at post-partum in all women with term birth. The number of *Lactobacillus* reads in each woman's microbiome did not increase as pregnancy progressed from 8 weeks of gestational age to 30 weeks of gestational age.

Glycogen is anaerobically converted to lactic acid in the vagina. *Lactobacilli* are the main component of lactic acid bacteria group. To determine whether the number of short reads predicted to participate in the metabolic activity in the *Lactobacillus* population genomes increase as pregnancy progresses and decrease following parturition, open reading frames (ORFs) were predicted in contigs participating in the *Lactobacillus* population genomes in each sample, and then aligned to the genes in the KEGG pathways [106]. The number of short reads predicted to participate in the metabolic activity in the *Lactobacillus* population genomes was counted and normalized in each sample. The number of short reads for predicted metabolic genes in the *Lactobacillus* population genomes was higher in the samples collected at 35–38 weeks of gestation than in the samples collected at post-partum in all women with term birth. It did not increase for any women as pregnancy progressed from 8 weeks of gestational age to 30 weeks of gestational age.

The conclusion from this study was that the vaginal microbiomes of women who delivered at term and who did not have a previous preterm birth history were generally dominated by *Lactobacillus* species during pregnancy. The vaginal microbiomes of women who delivered at term but had a previous preterm birth history were more diverse

than those of women who delivered at term and who did not have a previous preterm birth history.

5.2 Future research

In Chinese BV project, the level of sialidase activity and the relative abundances of *Prevotella* and *Gardnerella* species were detected to increase in all BV patients. However, we do not know whether *Prevotella* and *Gardnerella* species are the source of sialidase activity. Further study to isolate *Prevotella* and *Gardnerella* species from vaginal discharge of BV patients using culture-dependent methods and test their sialidase activity would help understand whether they are responsible for the increase of sialidase activity in BV patients.

In US BV project, I found that the bacteria *Aquabacterium*, *Chryseobacterium*, *Elizabethkingia*, *Finegoldia*, *Morganella*, *Negativicoccus*, *Phenylobacterium*, *Propionibacterium*, *Serratia*, and *Siphonobacter* were positively associated with the metabolite '2-Methyl-2-hydroxybutanoic acid' via the enzymes 'malate dehydrogenase' (K00024), '5-amino-6-(5-phosphoribosylamino) uracil reductase' (K11752), and 'undecaprenyl-diphosphatase' (K06153). Future studies to determine whether or how these ten bacteria produce the metabolite '2-Methyl-2-hydroxybutanoic acid' through these three functional genes are needed.

In pregnancy project, I found that the abundance of *Lactobacilli* and functional genes for metabolic activity in *Lactobacilli* did not significantly increase in women whose vaginal

microbiomes were dominated by *Lactobacilli* during pregnancy as pregnancy progressed from 8 weeks of gestational age to 30 weeks of gestational age. The increase may have happened before 8 weeks of gestational age. Future study to collect vaginal samples before 8 weeks of gestation during pregnancy would help understand how *Lactobacilli* change in very early pregnancy. Moreover, further studies to measure estrogen level, vaginal glycogen content, or vaginal pH value would help understand the direct relationship among estrogen level, vaginal glycogen content, colonization of *Lactobacillus* in vagina, and vaginal pH value.

CHAPTER 6. FIGURES

Figure 1. Shannon diversity indices of the vaginal microbiomes collected from women with and without bacterial vaginosis. The OTUs in the analysis are estimated based on 97% 16S rRNA sequence similarity. The average Shannon index was significantly different between women with and without BV, as identified by the Kruskal-Wallis rank sum test. P-values < 0.05 were regarded as significant. Error bar represents the standard error of the mean (SEM).

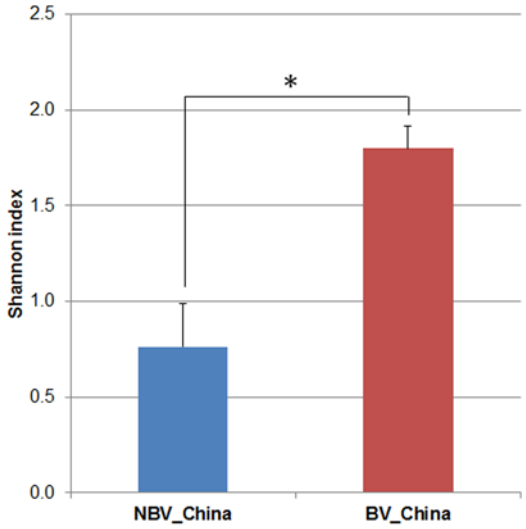


Figure 2. Bray-Curtis dissimilarity-based PCoA analysis showing clustering of vaginal microbiomes. Bray-Curtis dissimilarity-based PCoA was performed on OTUs from each vaginal sample. The OTUs in the analysis were estimated based on 97% 16S rRNA sequence similarity. Microbial communities that were significantly different between women with and without BV were identified by the PERMANOVA. P-values < 0.05 were regarded as significant. The oval in the figure represents that there is significant difference between two groups.

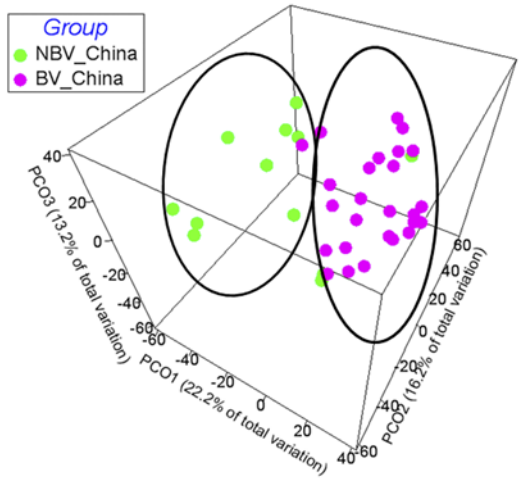


Figure 3. Selected KEGG orthologous groups whose relative abundances were higher in both Chinese BV patients compared to non-BV patients. The green color represents Chinese BV patients. The blue color represents Chinese women without BV.

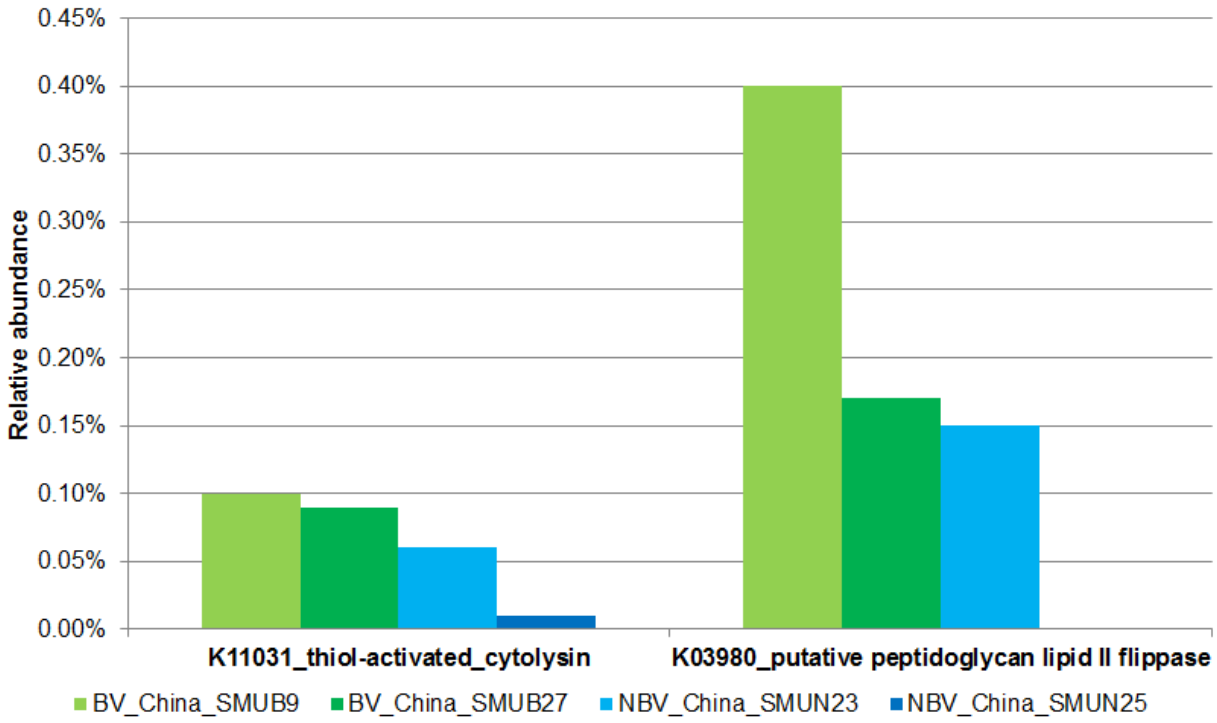


Figure 4. Bacterial genera more abundant in Chinese BV patients. Genera with significantly different abundances between Chinese women with and without BV were identified by the LEfSe. P-values < 0.05 were regarded as significant. The cell in the heatmap represents the relative percentage of each bacterium within each sample. The red color in the heatmap represents high abundance bacterial genera. The yellow color in the heatmap represents low abundance bacterial genera.

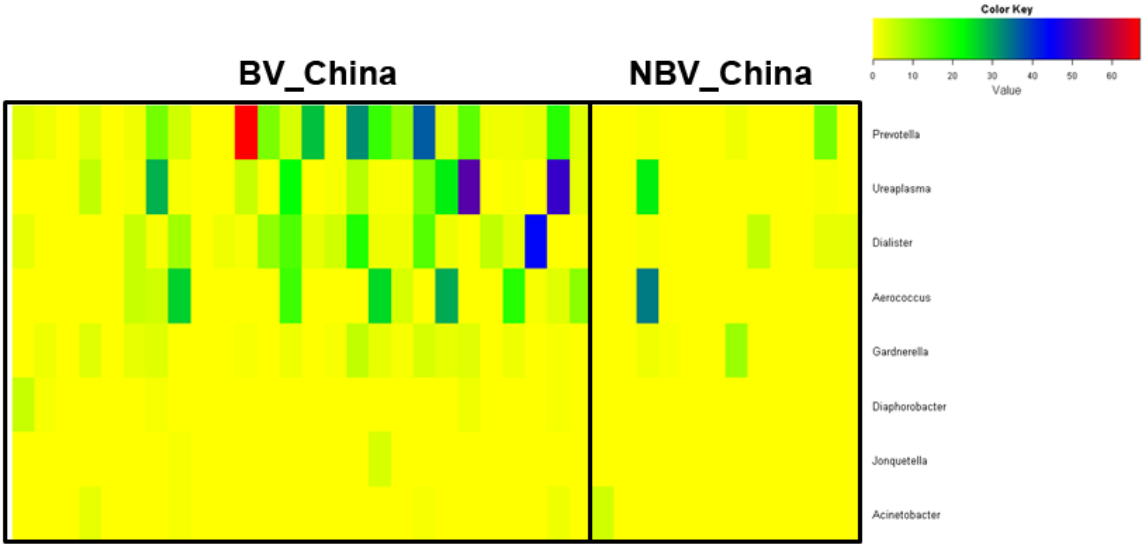


Figure 5. Selected bacterial strains whose relative abundances were higher in both Chinese BV patients. The green color represents Chinese BV patients. The blue color represents Chinese women without BV.

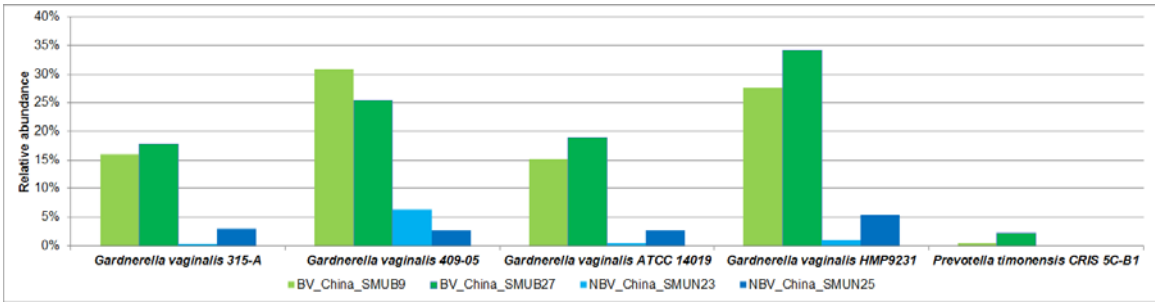


Figure 6. Comparative analysis of five reference genomes with assemblies from Chinese women with and without BV. (A) Comparative analysis of the *Gardnerella vaginalis 315-A* with the assemblies of Chinese samples, (B) Comparative analysis of the *Gardnerella vaginalis 409-05* with the assemblies of Chinese samples, (C) Comparative analysis of the *Gardnerella vaginalis ATCC-14019* with the assemblies of Chinese samples, (D) Comparative analysis of the *Gardnerella vaginalis HMP9231* with the assemblies of Chinese samples, (E) Comparative analysis of the *Prevotella timonensis CRIS 5C-B1* with the assemblies of Chinese samples. The inner purple ring represents the reference genome. Counting from inside, the first and second orange rings represent the assemblies from Chinese women with BV. The third and fourth green rings represent the assemblies from Chinese women without BV. The image was generated with GView server [126]. Samples that do not have assemblies against the reference genomes were not shown.

Figure 6 (cont.)

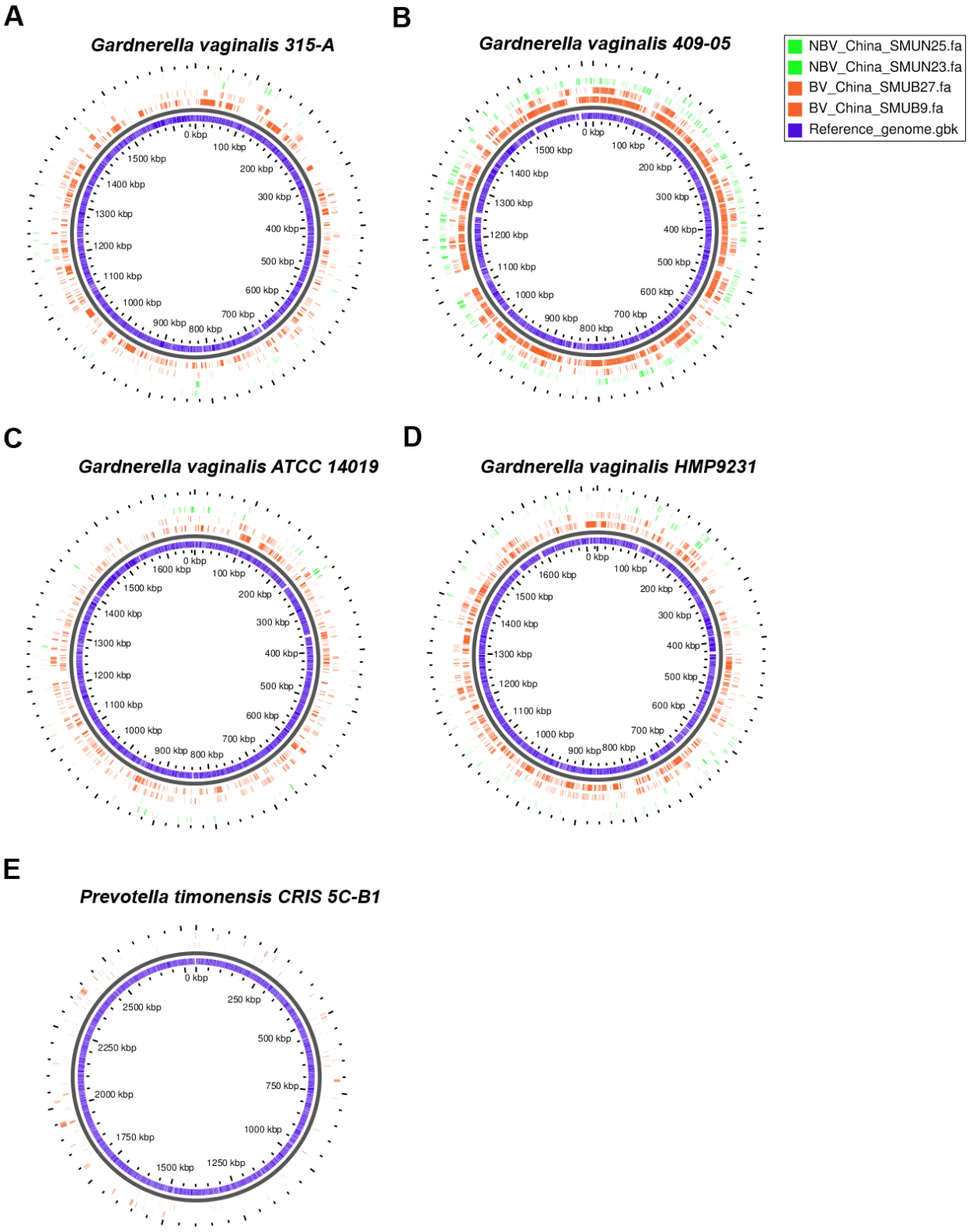


Figure 7. Overall network view of the correlations among bacterial genera, KEGG orthologous groups, and discharge-associated metabolites. Significant Spearman's rank correlation coefficients ($p < 0.05$) are shown as edges between genera/KEGG orthologous groups and genera/KEGG orthologous groups/metabolites. The p-values associated with these correlation coefficients are listed in Table 16. Pink nodes represent genera, yellow nodes KEGG orthologous groups, and green nodes metabolites. Red edges represent negative correlations; blue edges represent positive correlations.

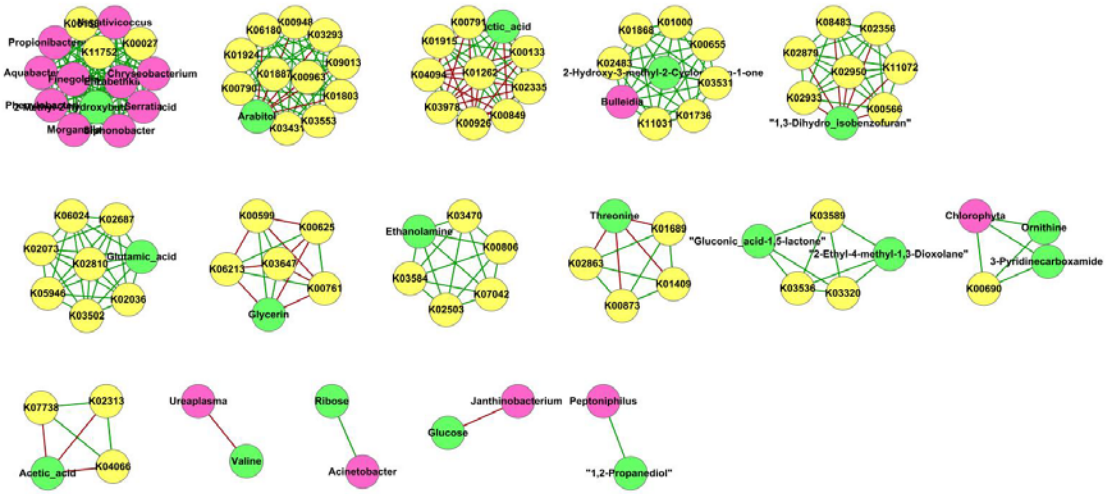


Figure 8. Sub-network view of the correlations between KEGG orthologous groups and discharge-associated metabolites. Significant Spearman's rank correlation coefficients ($p < 0.05$) are shown as edges between KEGG orthologous groups and KEGG orthologous groups/metabolites. The p-values associated with these correlation coefficients are listed in Table 16. Yellow nodes represent KEGG orthologous groups and green nodes metabolites. Red edges represent negative correlations; blue edges represent positive correlations.

Figure 8 (cont.)

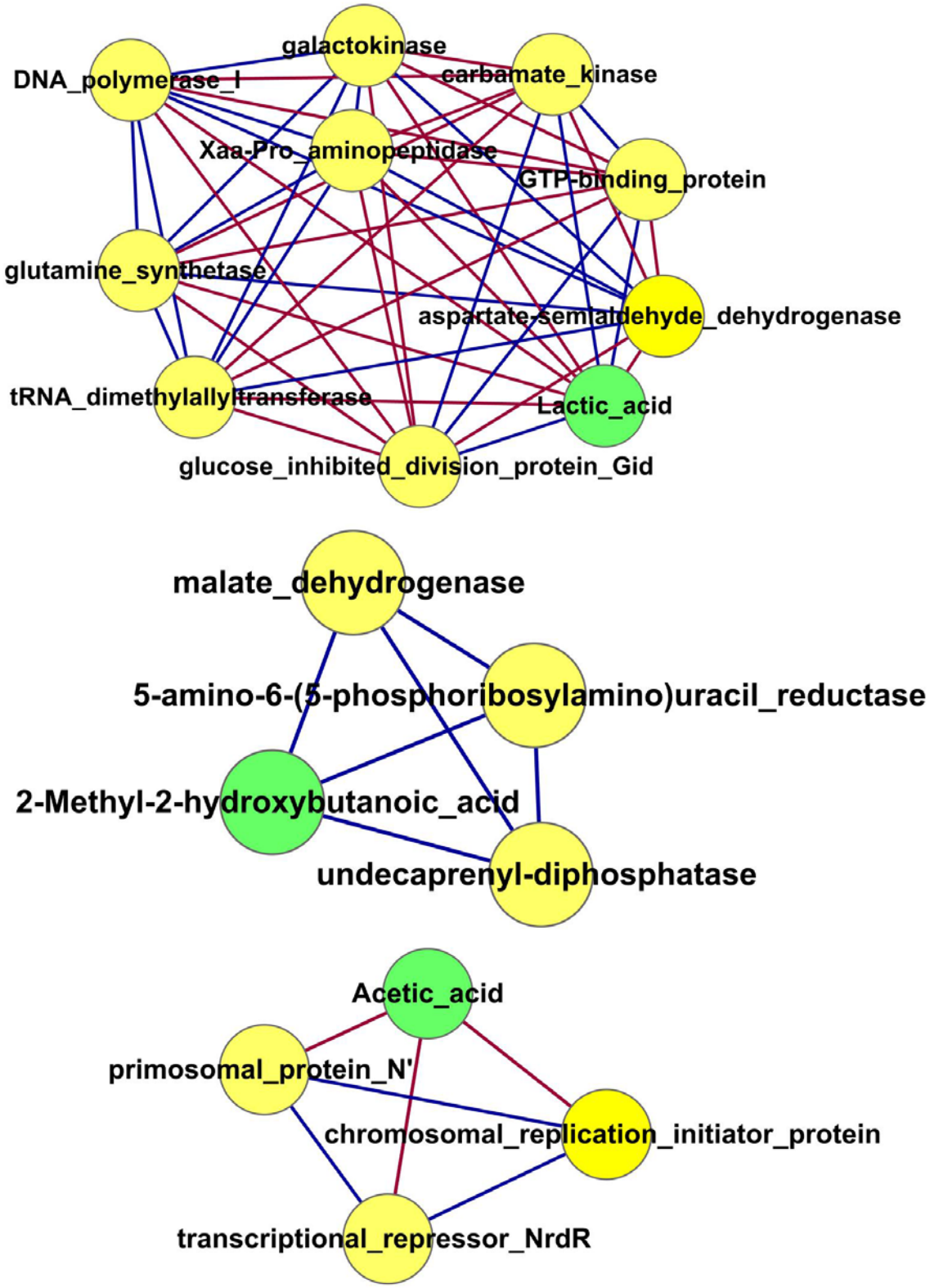


Figure 9. Sub-network view of the correlations between bacterial genera and discharge-associated metabolites. Significant Spearman's rank correlation coefficients ($p < 0.05$) are shown as edges between genera and genera/metabolites. The p-values associated with these correlation coefficients are listed in Table 16. Pink nodes represent genera and green nodes metabolites. Blue edges represent positive correlations.

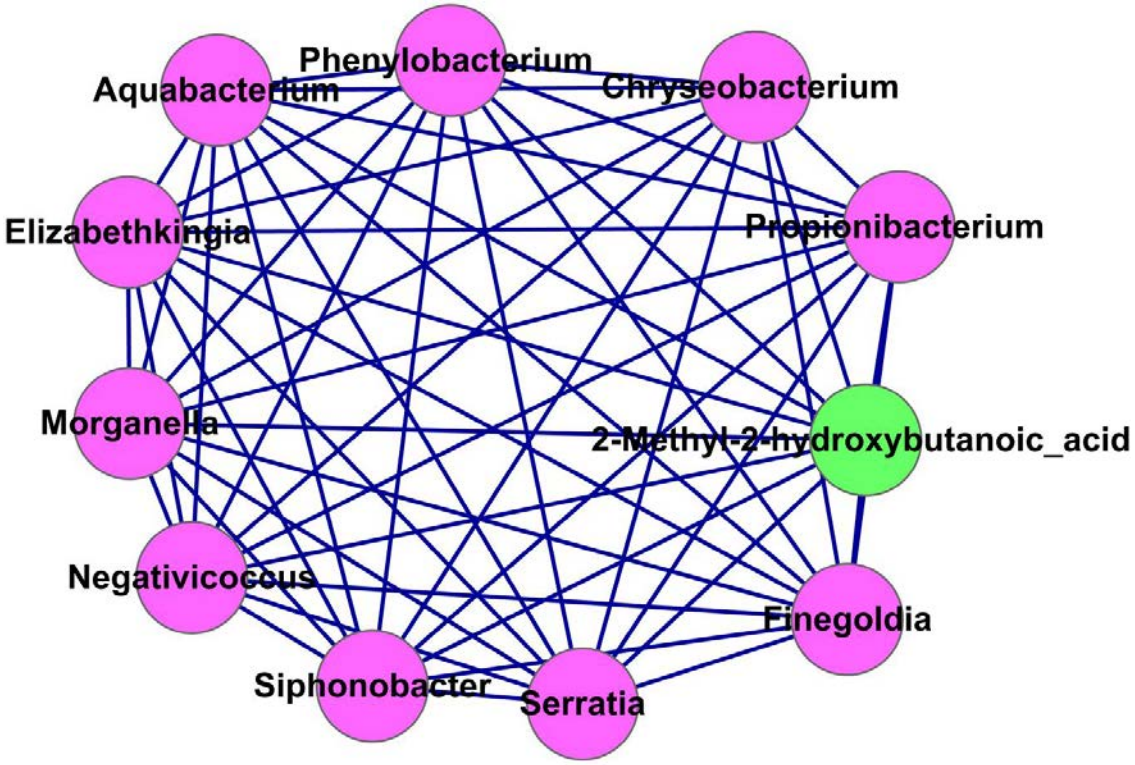


Figure 10. Sub-network view of the correlations between bacterial genera and KEGG orthologous groups. Significant Spearman's rank correlation coefficients ($p < 0.05$) are shown as edges between genera/KEGG orthologous groups and genera/KEGG orthologous groups. The p-values associated with these correlation coefficients are listed in Table 16. Pink nodes represent genera and yellow nodes KEGG orthologous groups. Blue edges represent positive correlations.

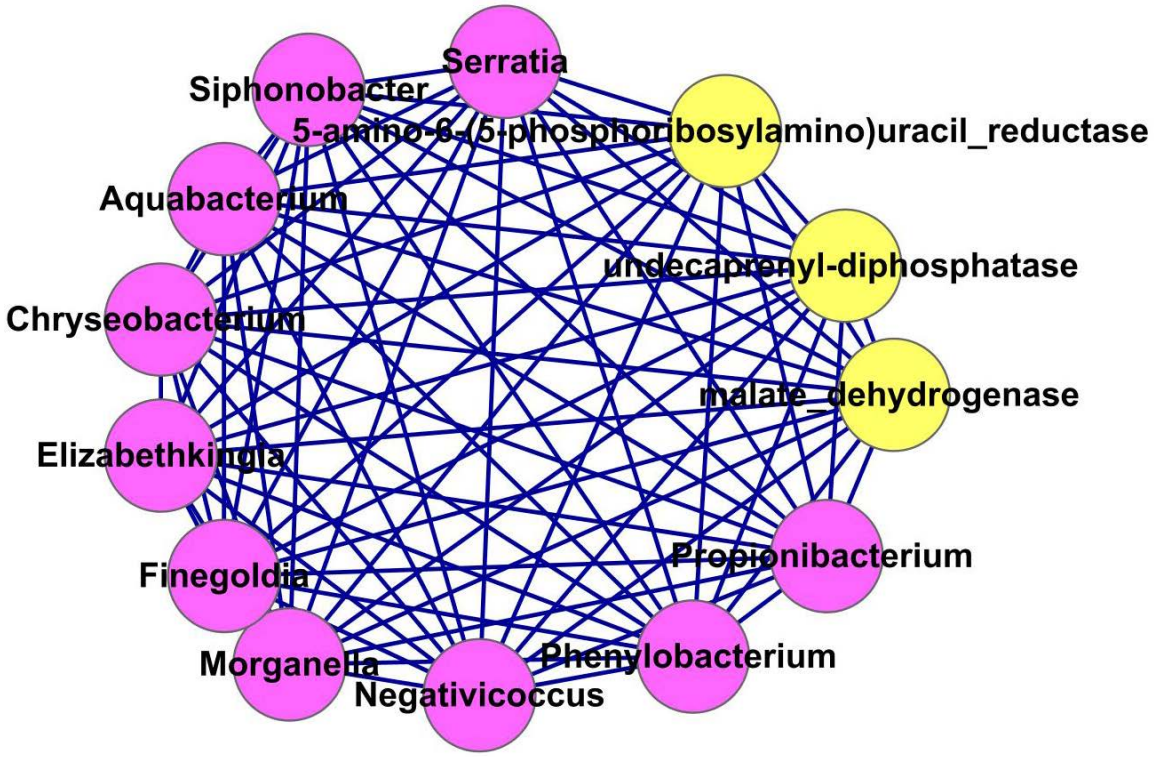


Figure 11. GroopM plots of twelve high-quality *Lactobacillus* population genomes. Dots represent contigs, colored by GC content. The size of each dot is proportional to the length of the contig. Subject 1XX represents women without a history of preterm birth. Subject 2XX represents women with a history of preterm birth.

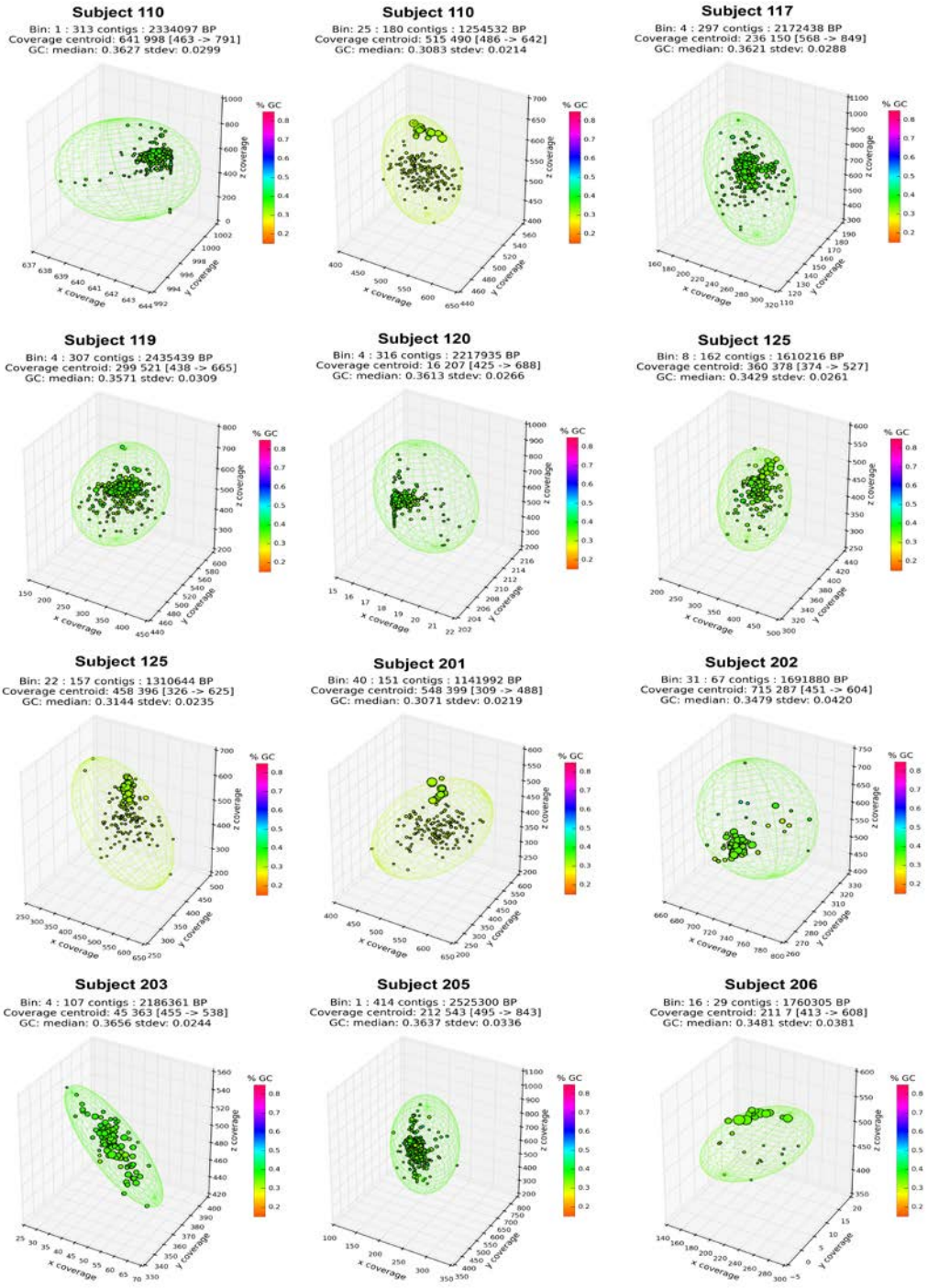


Figure 12. Heatmap of the relative abundance of bacterial species throughout pregnancy, during labor, and at post-partum. (A) The relative abundance of bacterial species was determined from the metagenomic short reads, (B) The relative abundance of bacterial species was determined from the metagenomic assembly. Each column represents one sample. Each row represents one bacterium. Only bacterial species at >5% abundance in at least one sample were shown in the heatmap. The red color represents high abundant bacterial species. The yellow color represents low abundant bacterial species. Samples in each subject are separated by vertical line. Women without a previous preterm birth history are named as low risk women. Women with a previous preterm birth history are named as high risk women.

Figure 12 (cont.)

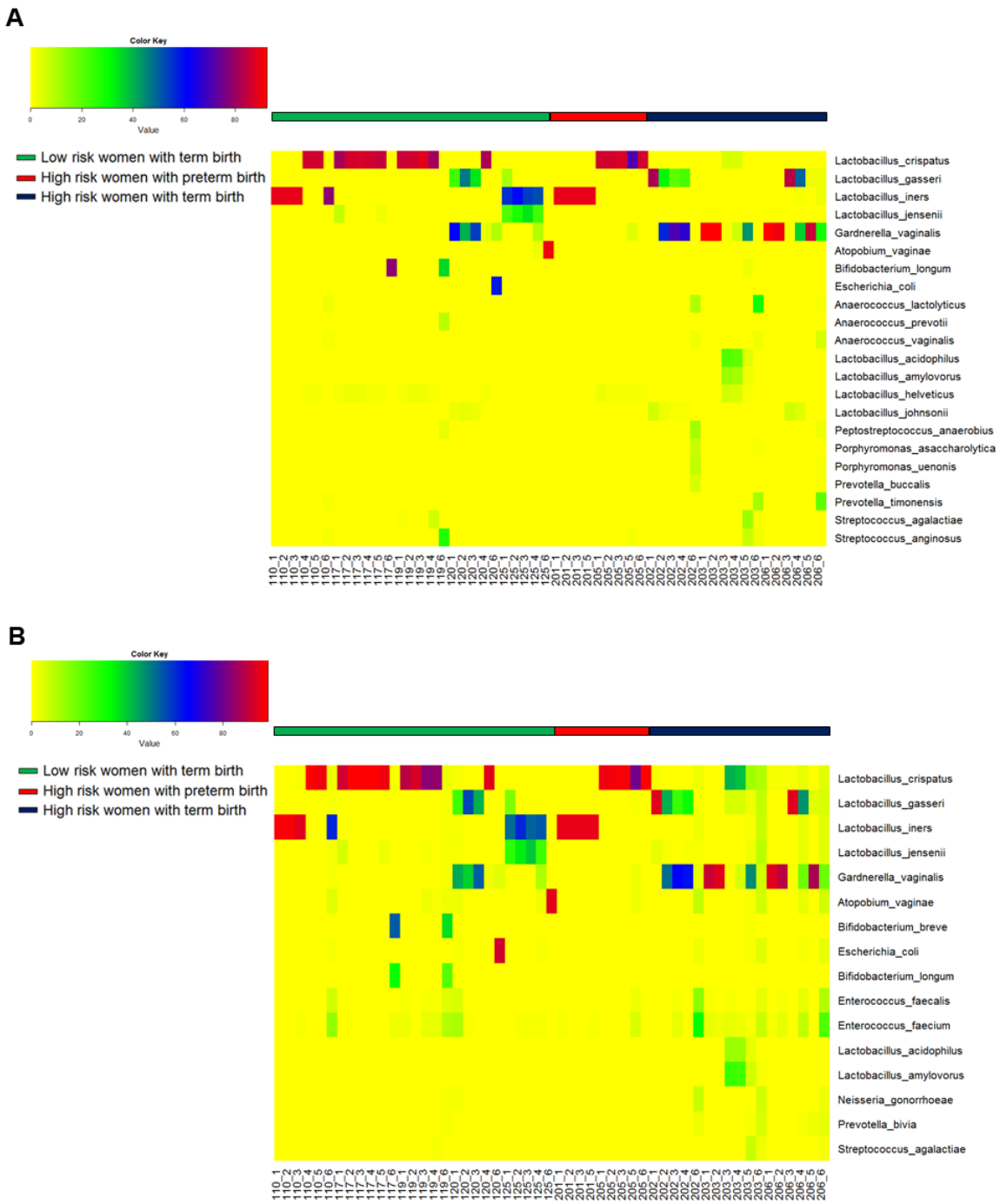


Figure 13. Majority vaginal bacteria throughout pregnancy, during labor, and at post-partum. (A) The majority vaginal bacteria were determined from the metagenomic short reads, (B) The majority vaginal bacteria was determined from the metagenomic assembly. Majority vaginal bacteria indicate that more than half of the bacteria in vagina. Others indicate that the majority vaginal bacteria are not any *Lactobacillus* species. NA represents that no sample is collected at this time point. Women without a previous preterm birth history are named as low risk women. Women with a previous preterm birth history are named as high risk women.

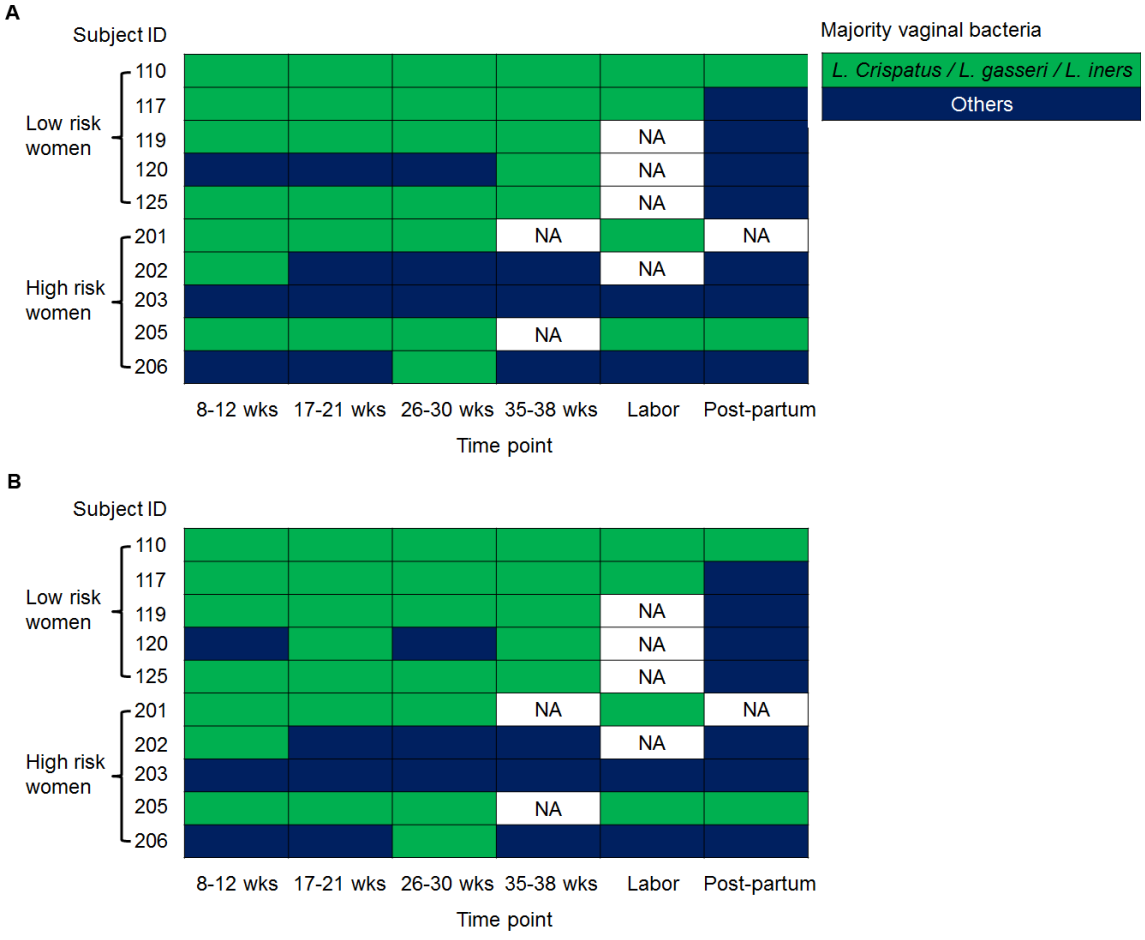


Figure 14. Relative abundance of *Lactobacillus spp.* in each subject at different time points. (A) The relative abundance of *Lactobacillus spp.* in women without a history of preterm birth determined from the metagenomic short reads, (B) The relative abundance of *Lactobacillus spp.* in women without a history of preterm birth determined from metagenomic assembly, (C) The relative abundance of *Lactobacillus spp.* in women with a history of preterm birth determined from the metagenomic short reads, (D) The relative abundance of *Lactobacillus spp.* in women with a history of preterm birth determined from metagenomic assembly. Subject 1XX represents women without a history of preterm birth. Subject 2XX represents women with a history of preterm birth. Subject 201 did not have samples collected at 35–38 weeks of gestation and post-partum. Subject 205 did not have samples collected at 35–38 weeks of gestation.

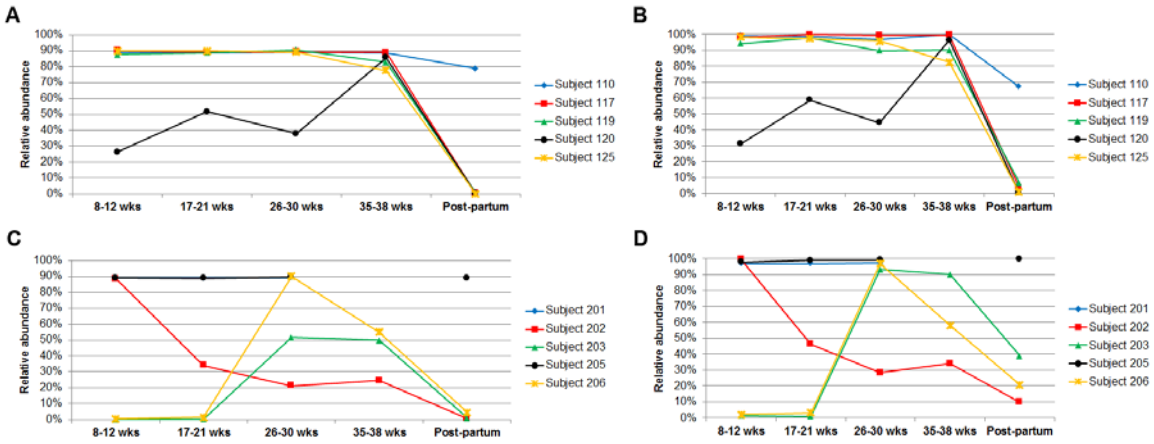


Figure 15. The number of short reads participating in the recovery of the *Lactobacillus* population genomes in each subject at different time points. (A) The number of short reads participating in the recovery of the *Lactobacillus* population genomes in women without a history of preterm birth, (B) The number of short reads participating in the recovery of the *Lactobacillus* population genomes in women with a history of preterm birth. Subject 1XX represents women without a history of preterm birth. Subject 2XX represents women with a history of preterm birth. Subject 201 did not have samples collected at 35–38 weeks of gestation and post-partum. Subject 205 did not have samples collected at 35–38 weeks of gestation. The number of short reads in y-axis is calculated based on the following formula: (number of short reads mapped to *Lactobacillus* population genomes / total number of clean reads) * 1,000,000.

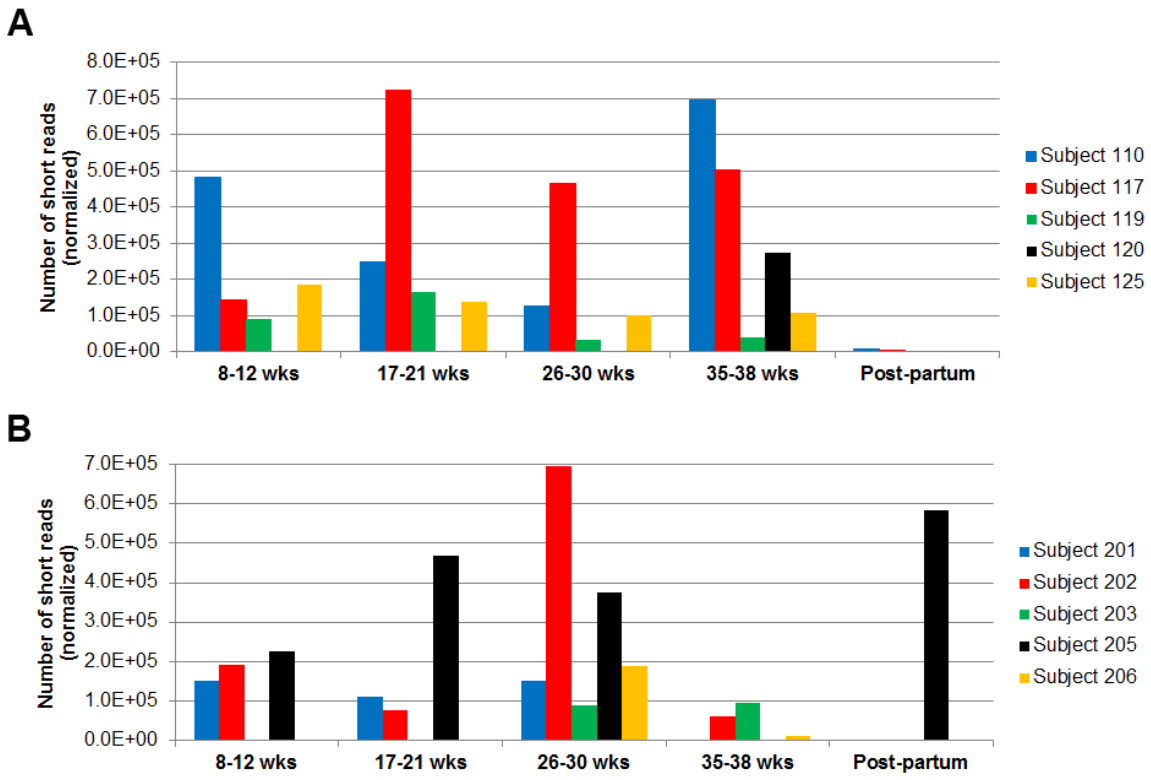
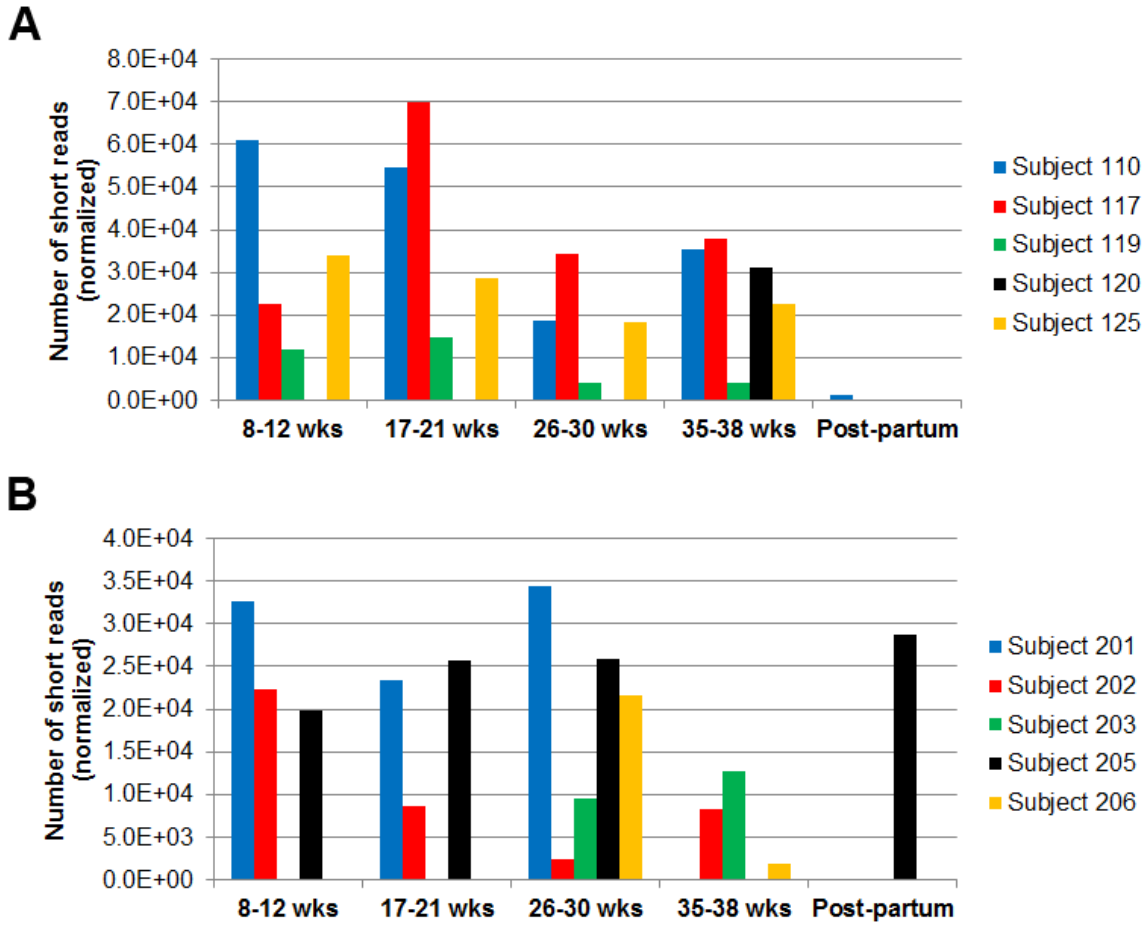


Figure 16. The number of short reads predicted to participate in the metabolic activity in the *Lactobacillus* population genomes in each subject at different time points. (A) The number of short reads predicted to participate in the metabolic activity in the *Lactobacillus* population genomes in women without a history of preterm birth, (B) the number of short reads predicted to participate in the metabolic activity in the *Lactobacillus* population genomes in women with a history of preterm birth. Subject 1XX represents women without a history of preterm birth. Subject 2XX represents women with a history of preterm birth. Subject 201 did not have samples collected at 35–38 weeks of gestation and post-partum. Subject 205 did not have samples collected at 35–38 weeks of gestation. The number of short reads in y-axis is calculated based on the following formula: (number of short reads predicted to participate in the metabolic activity in the *Lactobacillus* population genomes / total number of clean reads) * 100,000,000.

Figure 16 (cont.)



CHAPTER 7. TABLES

Table 1. Sample descriptions.

Sample ID	Disease Status	Race/Ethnicity	Sialidase Test	Age In Years	Length Of Residence	Marital Status	Sexual Orientation	Days Since Last Period	Tampon Use During Periods
SMUB7	BV	Han	Positive	38	38	Married	Heterosexual	5	Yes
SMUB8	BV	Han	Positive	25	25	Married	Heterosexual	15	Yes
SMUB9	BV	Han	Positive	24	24	Single	Heterosexual	8	Yes
SMUB19	BV	Han	Positive	36	36	Married	Heterosexual	3	Yes
SMUB23	BV	Han	Positive	32	32	Married	Heterosexual	6	Yes
SMUB25	BV	Han	Positive	36	36	Married	Heterosexual	3	Yes
SMUB26	BV	Han	Positive	26	26	Married	Heterosexual	10	Yes
SMUB27	BV	Han	Positive	47	47	Married	Heterosexual	20	Yes
SMUB29	BV	Han	Positive	42	42	Married	Heterosexual	17	Yes
SMUB30	BV	Han	Positive	29	29	Married	Heterosexual	6	Yes
SMUB31	BV	Han	Positive	20	20	Single	Heterosexual	5	Yes
SMUB37	BV	Han	Positive	34	34	Married	Heterosexual	9	Yes
SMUB74	BV	Han	Positive	43	43	Married	Heterosexual	11	Yes
SMUB75	BV	Han	Positive	37	37	Married	Heterosexual	11	Yes
SMUB77	BV	Han	Positive	26	26	Married	Heterosexual	5	Yes
SMUB78	BV	Han	Positive	38	38	Married	Heterosexual	8	Yes
SMUB79	BV	Han	Positive	37	37	Divorced	Heterosexual	18	Yes
SMUB81	BV	Han	Positive	35	35	Married	Heterosexual	4	Yes
SMUB82	BV	Han	Positive	43	43	Married	Heterosexual	3	Yes
SMUB83	BV	Han	Positive	50	50	Married	Heterosexual	20	Yes
SMUB84	BV	Han	Positive	37	37	Married	Heterosexual	16	Yes
SMUB85	BV	Han	Positive	46	46	Married	Heterosexual	24	Yes
SMUB86	BV	Han	Positive	40	40	Married	Heterosexual	10	Yes
SMUB87	BV	Han	Positive	38	38	Married	Heterosexual	20	Yes
SMUB88	BV	Han	Positive	31	31	Married	Heterosexual	6	Yes
SMUB89	BV	Han	Positive	44	44	Married	Heterosexual	3	Yes
SMUN5	NBV	Han	Negative	30	30	Married	Heterosexual	8	Yes
SMUN11	NBV	Han	Negative	33	33	Married	Heterosexual	19	Yes
SMUN12	NBV	Han	Negative	26	26	Married	Heterosexual	5	Yes
SMUN15	NBV	Han	Negative	35	35	Married	Heterosexual	6	Yes
SMUN20	NBV	Han	Negative	33	33	Married	Heterosexual	8	Yes
SMUN23	NBV	Han	Negative	32	32	Single	Heterosexual	12	Yes
SMUN25	NBV	Han	Negative	35	35	Married	Heterosexual	14	Yes
SMUN26	NBV	Han	Negative	34	34	Married	Heterosexual	3	Yes
SMUN27	NBV	Han	Negative	29	29	Married	Heterosexual	5	Yes
SMUN29	NBV	Han	Negative	52	52	Married	Heterosexual	Menopaus	Yes
SMUN31	NBV	Han	Negative	33	33	Divorced	Heterosexual	8	Yes
SMUN33	NBV	Han	Negative	48	48	Married	Heterosexual	16	Yes

Table 1 (cont.)

Sample ID	Number Of Previous Pregnancies	Number Of Sexual Partners In the Past 6	Gender of Sexual Partners In the Past 6	New Sexual Partner In the Past 6 Months	Days Since Intercourse	Number Of Vaginal Sex/Week	Number Of Oral Sex/Week	Number Of Anal Sex/Week	Condom Use	Number Of Bath/Week	Days Since Last Douche
SMUB7	2	1	Male	No	2	2	0	0	Yes	7	
SMUB8	0	1	Male	Yes	3	3	1	0	Yes	7	3
SMUB9	0	2	Male	Yes	1	4	0	0	Yes	7	
SMUB19	1	1	Male	No	3	2	0	0	No	7	
SMUB23	1	1	Male	No	6	3	0	0	No	7	
SMUB25	1	1	Male	No	2	3	0	0	No	7	
SMUB26	1	1	Male	No	3	2	0	0	Yes	7	
SMUB27	2	1	Male	No	1	2	0	0	No	7	2
SMUB29	1	1	Male	No	5	1	0	0	No	7	5
SMUB30	2	1	Male	No	2	2	1	0	No	7	
SMUB31	0	1	Male	Yes	2	3	0	0	Yes	7	
SMUB37	1	1	Male	No	10	1	0	0	No	7	
SMUB74	1	1	Male	No	2	2	0	0	No	7	
SMUB75	1	1	Male	No	4	2	0	0	No	7	
SMUB77	1	1	Male	No	3	3	0	0	Yes	7	3
SMUB78	2	1	Male	No	2	3	0	0	No	7	
SMUB79	3	1	Male	Yes	6	1	0	0	No	7	
SMUB81	1	0	Male	No	190	1	0	0	No	7	2
SMUB82	1	1	Male	No	3	2	0	0	No	7	
SMUB83	1	1	Male	No	30	1	0	0	No	7	
SMUB84	1	1	Male	No	2	3	0	0	No	7	3
SMUB85	1	1	Male	No	6	1	0	0	No	7	
SMUB86	1	1	Male	No	4	1	0	0	No	7	
SMUB87	2	1	Male	No	2	3	0	0	No	7	
SMUB88	2	1	Male	No	3	2	0	0	No	7	
SMUB89	3	1	Male	No	3	2	0	0	No	7	
SMUN5	1	1	Male	No	1	3	0	0	Yes	7	
SMUN11	1	1	Male	No	2	2	0	0	No	7	
SMUN12	1	1	Male	No	2	3	0	0	Yes	7	
SMUN15	4	1	Male	No	5	1	0	0	No	7	
SMUN20	1	1	Male	No	4	1	0	0	No	7	
SMUN23	1	1	Male	No	2	3	0	0	No	7	
SMUN25	1	1	Male	No	7	1	0	0	No	7	
SMUN26	1	1	Male	No	3	2	0	0	No	7	
SMUN27	1	1	Male	No	2	3	0	0	Yes	7	
SMUN29	1	0	Male	No	360	1	0	0	No	7	
SMUN31	1	1	Male	No	5	1	0	0	No	7	
SMUN33	1	1	Male	No	10	1	0	0	No	7	

Table 2. 16S rRNA sequence statistics of vaginal samples.

Sample ID	Total sequences number	Quality-filtered sequences number	Average quality-filtered sequence length (nt)
SMUB7	8847	6735	485
SMUB8	11327	9053	479
SMUB9	14521	10336	483
SMUB19	15049	11600	474
SMUB23	11364	8024	488
SMUB25	14503	9973	484
SMUB26	14862	10959	484
SMUB27	16205	12057	484
SMUB29	16397	12183	488
SMUB30	14624	10625	481
SMUB31	17130	10978	484
SMUB37	9654	6793	492
SMUB74	10978	7695	492
SMUB75	23701	17900	485
SMUB77	4657	3646	490
SMUB78	12131	9288	487
SMUB79	4322	3255	485
SMUB81	9387	6669	487
SMUB82	13074	8865	486
SMUB83	11046	7904	490
SMUB84	10454	8110	484
SMUB85	23757	18956	485
SMUB86	9647	7288	488
SMUB87	10905	8283	483
SMUB88	24229	17486	487
SMUB89	29088	19111	485
SMUN5	13469	9999	478
SMUN11	11294	6993	489
SMUN12	12286	9356	490
SMUN15	15392	12027	484
SMUN20	13088	7145	476
SMUN23	15406	10700	483
SMUN25	17523	13281	475
SMUN26	23210	18295	478
SMUN27	16216	12554	481
SMUN29	25617	19750	482
SMUN31	14241	11507	481
SMUN33	14287	9661	485

Table 3. List of Human Microbiome Project urogenital tract reference genomes.

Acinetobacter baumannii ATCC 19606 = CIP 70.34
Actinobaculum massiliae ACS-171-V-Col2
Actinomyces coleocanis DSM 15436
Actinomyces neuii BVS029A5
Actinomyces turicensis ACS-279-V-Col4
Actinomyces urogenitalis DSM 15434
Aerococcus urinae ACS-120-V-Col10a
Aerococcus viridans ATCC 11563
Anaerococcus hydrogenalis ACS-025-V-Sch4
Anaerococcus lactolyticus ATCC 51172
Anaerococcus prevotii ACS-065-V-Col13
Anaerococcus tetradius ATCC 35098
Atopobium vaginae DSM 15829
Atopobium vaginae PB189-T1-4
Bifidobacterium breve ACS-071-V-Sch8b
Bifidobacterium dentium ATCC 27679
Bifidobacterium dentium JCVIHMP022
Brevibacterium mcbrellneri ATCC 49030
Chryseobacterium gleum ATCC 35910
Clostridiales genomosp. BVAB3 str. UPII9-5
Corynebacterium aurimucosum ATCC 700975
Corynebacterium genitalium ATCC 33030
Corynebacterium glucuronolyticum ATCC 51866
Corynebacterium glucuronolyticum ATCC 51867
Corynebacterium jeikeium ATCC 43734
Corynebacterium lipophiloflavum DSM 44291
Corynebacterium pseudogenitalium ATCC 33035
Corynebacterium striatum ATCC 6940
Dialister micraerophilus DSM 19965
Dialister microaerophilus UPII 345-E
Enterococcus durans FB129-CNAB-4
Enterococcus faecalis ATCC 29200
Enterococcus faecalis ERV81
Enterococcus faecalis ERV85
Enterococcus faecalis ERV93
Enterococcus faecalis HH22
Enterococcus faecalis TX0312
Enterococcus faecalis TX0635
Enterococcus faecalis TX0855

Table 3. (cont.)

Enterococcus faecium E422
Enterococcus faecium P1139
Enterococcus faecium S447
Eremococcus coleocola ACS-139-V-Col8
Escherichia coli 83972
Finegoldia magna ACS-171-V-Col3
Finegoldia magna ATCC 53516
Finegoldia magna BVS033A4
Finegoldia magna SY403409CC001050417
Fusobacterium nucleatum subsp. nucleatum ATCC 2372
Gardnerella vaginalis 315-A
Gardnerella vaginalis 409-05
Gardnerella vaginalis ATCC 14019
Gardnerella vaginalis HMP9231
Haemophilus parainfluenzae HK262
Klebsiella oxytoca 10-5245
Klebsiella oxytoca 10-5250
Lactobacillus coleohominis 101-4-CHN
Lactobacillus crispatus 125-2-CHN
Lactobacillus crispatus 214-1
Lactobacillus crispatus CTV-05
Lactobacillus crispatus FB049-03
Lactobacillus crispatus FB077-07
Lactobacillus crispatus JV-V01
Lactobacillus crispatus MV-1A-US
Lactobacillus crispatus MV-3A-US
Lactobacillus crispatus SJ-3C-US
Lactobacillus delbrueckii subsp. bulgaricus PB2003/044-
Lactobacillus fermentum 28-3-CHN
Lactobacillus gasseri 202-4
Lactobacillus gasseri 224-1
Lactobacillus gasseri JV-V03
Lactobacillus gasseri MV-22
Lactobacillus gasseri SJ-9E-US
Lactobacillus gasseri SV-16A-US
Lactobacillus iners ATCC 55195
Lactobacillus iners DSM 13335
Lactobacillus iners LactinV 01V1-a
Lactobacillus iners LactinV 03V1-b

Table 3. (cont.)

Lactobacillus iners LactinV 09V1-c
Lactobacillus iners LactinV 11V1-d
Lactobacillus iners LEAF 2052A-d
Lactobacillus iners LEAF 2053A-b
Lactobacillus iners LEAF 2062A-h1
Lactobacillus iners LEAF 3008A-a
Lactobacillus iners SPIN 1401G
Lactobacillus iners SPIN 2503V10-D
Lactobacillus iners UPII 143-D
Lactobacillus iners UPII 60-B
Lactobacillus jensenii 1153
Lactobacillus jensenii 115-3-CHN
Lactobacillus jensenii 269-3
Lactobacillus jensenii 27-2-CHN
Lactobacillus jensenii JV-V16
Lactobacillus jensenii SJ-7A-US
Lactobacillus johnsonii ATCC 33200
Lactobacillus oris PB013-T2-3
Lactobacillus salivarius ACS-116-V-Col5a
Lactobacillus vaginalis ATCC 49540
Megasphaera genomosp. type_1 str. 28L
Megasphaera sp. UPII 135-E
Megasphaera sp. UPII 199-6
Mobiluncus curtisii ATCC 43063
Mobiluncus curtisii ATCC 51333
Mobiluncus curtisii subsp. curtisii ATCC 35241
Mobiluncus curtisii subsp. holmesii ATCC 35242
Mobiluncus mulieris 28-1
Mobiluncus mulieris ATCC 35239
Mobiluncus mulieris ATCC 35243
Mobiluncus mulieris FB024-16
Mycobacterium parascrofulaceum ATCC BAA-614
Myroides odoratimimus CCUG 12901
Myroides odoratimimus CCUG 3837
Pasteurella bettyae CCUG 2042
Peptoniphilus duerdenii ATCC BAA-1640
Peptoniphilus harei ACS-146-V-Sch2b
Peptoniphilus lacrimalis 315-B
Porphyromonas asaccharolytica PR426713P-I

Table 3. (cont.)

Porphyromonas uenonis 60-3
Prevotella amnii CRIS 21A-A
Prevotella bivia JCVIHMP010
Prevotella buccalis ATCC 35310
Prevotella denticola CRIS 18C-A
Prevotella disiens FB035-09AN
Prevotella oralis ATCC 33269
Prevotella timonensis CRIS 5C-B1
Propionibacterium sp. 409-HC1
Propionibacterium sp. 434-HC2
Proteus mirabilis ATCC 29906
Roseomonas cervicalis ATCC 49957
Sphingobacterium spiritivorum ATCC 33300
Sphingobacterium spiritivorum ATCC 33861
Staphylococcus aureus subsp. aureus MN8
Staphylococcus epidermidis BVS058A4
Staphylococcus lugdunensis ACS-027-V-Sch2
Staphylococcus simulans ACS-120-V-Sch1
Streptococcus bovis ATCC 700338
Streptococcus pseudoporcinus SPIN 20026
Streptococcus urinalis FB127-CNA-2
Treponema phagedenis F0421
Veillonella atypica ACS-049-V-Sch6
Veillonella atypica ACS-134-V-Col7a
Veillonella parvula ACS-068-V-Sch12
Veillonella ratti ACS-216-V-Col6b

Table 4. Summary of the percentage of bases covered for each HMP urogenital tract reference genome in each sample.

	BV_SMUB9	BV_SMUB27	NBV_SMUN23	NBV_SMUN25
Acinetobacter_baumannii_ATCC_19606_=_CIP_70.34	0.00%	0.08%	0.01%	0.13%
Actinobaculum_massiliae_ACS-171-V-Col2	0.00%	0.00%	0.00%	0.00%
Actinomyces_coleocanis_DSM_15436	0.00%	0.00%	0.00%	0.00%
Actinomyces_neuii_BVS029A5	0.00%	0.16%	0.00%	0.00%
Actinomyces_turicensis_ACS-279-V-Col4	0.00%	0.00%	0.00%	0.00%
Actinomyces_urogenitalis_DSM_15434	0.00%	0.00%	0.00%	0.00%
Aerococcus_urinae_ACS-120-V-Col10a	0.02%	0.00%	0.00%	0.00%
Aerococcus_viridans_ATCC_11563	0.00%	0.00%	0.00%	0.00%
Anaerococcus_hydrogenalis_ACS-025-V-Sch4	2.10%	0.73%	0.02%	0.01%
Anaerococcus_lactolyticus_ATCC_51172	1.22%	0.00%	0.00%	0.00%
Anaerococcus_prevotii_ACS-065-V-Col13	0.08%	0.01%	0.00%	0.00%
Anaerococcus_tetradius_ATCC_35098	0.12%	0.01%	0.00%	0.00%
Atopobium_vaginae_DSM_15829	87.97%	0.25%	0.47%	0.52%
Atopobium_vaginae_PB189-T1-4	0.72%	0.00%	0.04%	0.00%
Bifidobacterium_breve_ACS-071-V-Sch8b	0.00%	0.00%	0.00%	0.00%
Bifidobacterium_dentium_ATCC_27679	0.00%	0.00%	0.00%	0.00%
Bifidobacterium_dentium_JCVIHMP022	0.00%	0.00%	0.03%	0.02%
Brevibacterium_mcbrellneri_ATCC_49030	0.00%	0.00%	0.00%	0.00%
Chryseobacterium_gleum_ATCC_35910	0.00%	0.00%	0.38%	0.00%
Clostridiales_genomosp._BVAB3_str._UPII9-5	0.26%	0.10%	0.05%	0.05%
Corynebacterium_aurimucosum_ATCC_700975	0.00%	0.00%	0.00%	0.00%
Corynebacterium_genitalium_ATCC_33030	0.00%	0.00%	0.00%	0.00%
Corynebacterium_glucuronolyticum_ATCC_51866	0.01%	0.00%	0.01%	0.00%
Corynebacterium_glucuronolyticum_ATCC_51867	0.00%	0.00%	0.00%	0.00%
Corynebacterium_jeikeium_ATCC_43734	0.00%	0.00%	0.00%	0.00%
Corynebacterium_lipophiloflavum_DSM_44291	0.00%	0.00%	0.00%	0.01%
Corynebacterium_pseudogenitalium_ATCC_33035	0.01%	0.00%	0.00%	0.00%
Corynebacterium_striatum_ATCC_6940	0.00%	0.00%	0.01%	0.01%
Dialister_micraerophilus_DSM_19965	0.88%	0.24%	0.15%	0.00%
Dialister_microaerophilus_UPII_345-E	0.85%	0.06%	0.05%	0.00%
Enterococcus_durans_FB129-CNAB-4	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_ATCC_29200	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_ERV81	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_ERV85	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_ERV93	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_HH22	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_TX0312	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_TX0635	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_TX0855	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecium_E422	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecium_P1139	0.01%	0.00%	0.00%	0.00%
Enterococcus_faecium_S447	0.00%	0.00%	0.00%	0.00%
Eremococcus_coleocola_ACS-139-V-Col8	0.03%	0.00%	0.00%	0.00%
Escherichia_coli_83972	0.00%	0.00%	0.36%	0.33%
Finegoldia_magna_ACS-171-V-Col3	0.94%	0.07%	0.00%	0.00%
Finegoldia_magna_ATCC_53516	0.07%	0.03%	0.00%	0.00%

Table 4. (cont.)

Finegoldia_magna_BVS033A4	0.71%	0.07%	0.00%	0.00%
Finegoldia_magna_SY403409CC001050417	0.70%	0.05%	0.01%	0.00%
Fusobacterium_nucleatum_subsp._nucleatum_ATCC_23726	0.01%	0.00%	0.00%	0.00%
Gardnerella_vaginalis_315-A	10.76%	2.91%	0.40%	1.04%
Gardnerella_vaginalis_409-05	74.44%	56.97%	12.14%	17.02%
Gardnerella_vaginalis_ATCC_14019	7.48%	1.75%	0.44%	0.39%
Gardnerella_vaginalis_HMP9231	7.36%	1.66%	0.21%	1.82%
Haemophilus_parainfluenzae_HK262	0.00%	0.00%	0.00%	0.00%
Klebsiella_oxytoca_10-5245	0.00%	0.00%	0.01%	0.00%
Klebsiella_oxytoca_10-5250	0.00%	0.00%	0.00%	0.00%
Lactobacillus_coleohominis_101-4-CHN	0.00%	0.00%	0.00%	0.00%
Lactobacillus_crispatus_125-2-CHN	0.00%	0.00%	0.01%	0.01%
Lactobacillus_crispatus_214-1	0.00%	0.00%	0.00%	0.02%
Lactobacillus_crispatus_CTV-05	0.01%	0.00%	0.00%	0.00%
Lactobacillus_crispatus_FB049-03	0.00%	0.00%	0.00%	0.00%
Lactobacillus_crispatus_FB077-07	0.01%	0.00%	0.01%	0.01%
Lactobacillus_crispatus_JV-V01	0.00%	0.00%	0.00%	0.00%
Lactobacillus_crispatus_MV-1A-US	0.00%	0.00%	0.00%	0.00%
Lactobacillus_crispatus_MV-3A-US	0.00%	0.00%	0.00%	0.00%
Lactobacillus_crispatus_SJ-3C-US	0.00%	0.00%	0.00%	0.00%
Lactobacillus_delbrueckii_subsp._bulgaricus_PB2003/044-T3	0.00%	0.00%	0.00%	0.00%
Lactobacillus_fermentum_28-3-CHN	0.00%	0.00%	0.00%	0.00%
Lactobacillus_gasseri_202-4	0.00%	0.01%	0.00%	0.04%
Lactobacillus_gasseri_224-1	0.05%	0.02%	0.11%	0.11%
Lactobacillus_gasseri_JV-V03	0.00%	0.01%	0.01%	0.01%
Lactobacillus_gasseri_MV-22	0.00%	0.00%	0.00%	0.00%
Lactobacillus_gasseri_SJ-9E-US	0.00%	0.00%	0.00%	0.00%
Lactobacillus_gasseri_SV-16A-US	0.00%	0.00%	0.00%	0.00%
Lactobacillus_iners_ATCC_55195	0.11%	0.24%	0.03%	0.55%
Lactobacillus_iners_DSM_13335	0.10%	0.14%	0.00%	0.72%
Lactobacillus_iners_LactinV_01V1-a	0.07%	0.03%	0.00%	0.60%
Lactobacillus_iners_LactinV_03V1-b	0.21%	0.21%	0.00%	1.04%
Lactobacillus_iners_LactinV_09V1-c	0.15%	0.11%	0.00%	0.55%
Lactobacillus_iners_LactinV_11V1-d	0.07%	0.03%	0.00%	0.43%
Lactobacillus_iners_LEAF_2052A-d	0.11%	0.15%	0.00%	0.41%
Lactobacillus_iners_LEAF_2053A-b	0.26%	0.21%	0.13%	0.86%
Lactobacillus_iners_LEAF_2062A-h1	0.08%	0.08%	0.00%	0.30%
Lactobacillus_iners_LEAF_3008A-a	0.10%	0.06%	0.00%	0.52%
Lactobacillus_iners_SPIN_1401G	0.12%	0.12%	0.01%	1.54%
Lactobacillus_iners_SPIN_2503V10-D	0.07%	0.01%	0.00%	0.22%
Lactobacillus_iners_UPII_143-D	0.03%	0.10%	0.00%	0.17%
Lactobacillus_iners_UPII_60-B	0.22%	0.40%	0.00%	0.41%
Lactobacillus_jensenii_1153	0.01%	0.00%	0.01%	0.00%
Lactobacillus_jensenii_115-3-CHN	0.00%	0.00%	0.00%	0.00%
Lactobacillus_jensenii_269-3	0.00%	0.00%	0.00%	0.00%
Lactobacillus_jensenii_27-2-CHN	0.00%	0.00%	0.00%	0.00%
Lactobacillus_jensenii_JV-V16	0.00%	0.01%	0.01%	0.00%
Lactobacillus_jensenii_SJ-7A-US	0.01%	0.10%	0.02%	0.01%
Lactobacillus_johnsonii_ATCC_33200	0.00%	0.00%	0.00%	0.00%

Table 4. (cont.)

Lactobacillus_oris_PB013-T2-3	0.00%	0.00%	0.00%	0.00%
Lactobacillus_salivarius_ACS-116-V-Col5a	0.00%	0.00%	0.00%	0.00%
Lactobacillus_vaginalis_ATCC_49540	0.01%	0.01%	0.06%	0.05%
Megasphaera_genomosp._type_1_str._28L	0.02%	0.00%	0.00%	0.06%
Megasphaera_sp._UPII_135-E	0.10%	0.01%	0.00%	0.04%
Megasphaera_sp._UPII_199-6	0.04%	0.01%	0.00%	0.00%
Mobiluncus_curtisii_ATCC_43063	0.29%	0.00%	0.15%	0.02%
Mobiluncus_curtisii_ATCC_51333	0.34%	0.00%	0.14%	0.01%
Mobiluncus_curtisii_subsp._curtisii_ATCC_35241	0.01%	0.00%	0.00%	0.00%
Mobiluncus_curtisii_subsp._holmesii_ATCC_35242	0.01%	0.00%	0.00%	0.00%
Mobiluncus_mulieris_28-1	0.22%	0.00%	0.12%	0.09%
Mobiluncus_mulieris_ATCC_35239	0.01%	0.00%	0.00%	0.00%
Mobiluncus_mulieris_ATCC_35243	0.00%	0.00%	0.00%	0.00%
Mobiluncus_mulieris_FB024-16	0.00%	0.00%	0.00%	0.00%
Mycobacterium_parascrofulaceum_ATCC_BAA-614	0.00%	0.00%	0.00%	0.00%
Myroides_odoratimimus_CCUG_12901	0.00%	0.00%	0.00%	0.00%
Myroides_odoratimimus_CCUG_3837	0.00%	0.00%	0.00%	0.00%
Pasteurella_bettyae_CCUG_2042	0.00%	0.05%	0.00%	0.00%
Peptoniphilus_duerdenii_ATCC_BAA-1640	0.32%	0.04%	0.01%	0.00%
Peptoniphilus_harei_ACS-146-V-Sch2b	12.40%	0.85%	0.01%	0.01%
Peptoniphilus_lacrimalis_315-B	0.22%	0.30%	0.00%	0.01%
Porphyromonas_asaccharolytica_PR426713P-I	0.01%	0.01%	0.00%	0.00%
Porphyromonas_uenonis_60-3	0.03%	0.00%	0.00%	0.00%
Prevotella_amnii_CRIS_21A-A	0.49%	0.20%	0.04%	0.06%
Prevotella_bivia_JCVIHMP010	25.14%	1.77%	0.10%	0.03%
Prevotella_buccalis_ATCC_35310	2.44%	1.16%	0.01%	0.00%
Prevotella_denticola_CRIS_18C-A	0.28%	0.15%	0.00%	0.00%
Prevotella_disiens_FB035-09AN	18.30%	0.76%	0.01%	0.01%
Prevotella_oralis_ATCC_33269	0.08%	0.00%	0.00%	0.00%
Prevotella_timonensis_CRIS_5C-B1	32.44%	23.95%	0.03%	0.01%
Propionibacterium_sp._409-HC1	0.01%	0.00%	0.00%	0.00%
Propionibacterium_sp._434-HC2	0.00%	0.00%	0.00%	0.00%
Proteus_mirabilis_ATCC_29906	0.00%	0.00%	0.00%	0.00%
Roseomonas_cervicalis_ATCC_49957	0.00%	0.00%	0.00%	0.00%
Sphingobacterium_spiritivorum_ATCC_33300	0.00%	0.00%	0.00%	0.00%
Sphingobacterium_spiritivorum_ATCC_33861	0.00%	0.00%	0.10%	0.00%
Staphylococcus_aureus_subsp._aureus_MN8	0.00%	0.01%	0.00%	0.00%
Staphylococcus_epidermidis_BVS058A4	0.00%	0.01%	0.06%	0.01%
Staphylococcus_lugdunensis_ACS-027-V-Sch2	0.00%	0.00%	0.01%	0.00%
Staphylococcus_simulans_ACS-120-V-Sch1	0.00%	0.00%	0.02%	0.02%
Streptococcus_bovis_ATCC_700338	0.00%	0.01%	0.00%	0.00%
Streptococcus_pseudoporcinus_SPIN_20026	0.07%	0.01%	0.00%	0.00%
Streptococcus_urinalis_FB127-CNA-2	0.12%	0.01%	0.02%	0.00%
Treponema_phagedenis_F0421	0.00%	0.00%	0.00%	0.00%
Veillonella_atypica_ACS-049-V-Sch6	0.00%	0.00%	0.00%	0.00%
Veillonella_atypica_ACS-134-V-Col7a	0.00%	0.00%	0.01%	0.00%
Veillonella_parvula_ACS-068-V-Sch12	0.00%	0.00%	0.00%	0.00%
Veillonella_ratti_ACS-216-V-Col6b	0.28%	0.00%	0.00%	0.00%
Note: only reads (mapping quality > 20) were considered as correct alignment.				

Table 5. List of scaffolds whose sequencing depths were bigger than 20X in Chinese BV patients and which were classified by essential single copy gene.

Taxonomic ID	Scaffold ID	Scaffold length	GC (%)	Sequencing depth in subject SMUB9	Sequencing depth in subject SMUB27
Actinobacteria	11	9955	38.37	24.684X	20.748X
Actinobacteria	31	16965	37.96	25.598X	21.25X
Actinobacteria	103	19418	39.78	26.577X	25.138X
Actinobacteria	105	13609	39.38	30.149X	26.823X
Actinobacteria	3659	5170	37.81	25.955X	24.831X
Actinobacteria	3665	12617	39.72	26.426X	26.474X
Actinobacteria	12	3365	43.42	32.929X	30.281X
Actinobacteria	26	4827	39.24	33.123X	20.731X
Actinobacteria	40	21633	38.39	29.12X	23.425X
Actinobacteria	68	10933	38.93	27.66X	25.003X
Actinobacteria	81	7564	37.36	21.597X	24.519X
Actinobacteria	104	17692	39.72	27.611X	26.973X
Actinobacteria	144	9311	39.3	27.43X	24.95X
Actinobacteria	1987	10015	42.19	40.205X	27.334X

Note: Reads were independently mapped from subjects SMUB9 and SMUB27 to the assembled scaffolds from SMUB9 or SMUB27 to calculate sequencing depth for each subject. Only scaffolds (sequencing depth > 20 in both subjects) are listed in the table.

Table 6. Sequence statistics of metagenome vaginal samples.

Sample ID	SMUB9	SMUB27	SMUN23	SMUN25
Number of bases in raw metagenome	2,653,178,696	2,738,728,726	2,602,034,922	1,596,228,644
Number of bases in quality-filtered metagenome	409,491,561	510,629,843	568,325,264	208,244,881
Number of contigs in assembly #1	2,031	718	101	31
Total size of contigs in assembly #1 (bp)	2,286,994	490,128	49,770	14,449
Longest contig length in assembly #1 (bp)	36,300	5,656	1,986	1,431
N50 contig length in assembly #1 (bp)	1,726	722	472	424
Number of bases aligned to assembly #1	38,693,348	5,149,222	898,779	115,612
Number of contigs in assembly #2	129	157	0	0
Total size of contigs in assembly #2 (bp)	174,885	164,866	0	0
Longest contig length in assembly #2 (bp)	15,092	14,930	0	0
N50 contig length in assembly #2 (bp)	2,128	1,669	0	0
Number of bases aligned to assembly #2	3,238,746	2,512,421	0	0
Number of contigs in assembly #3	3,793	2,010	6,067	1,856
Total size of contigs in assembly #3 (bp)	3,945,182	1,701,892	3,303,190	898,952
Longest contig length in assembly #3 (bp)	22,569	21,718	13,528	6,803
N50 contig length in assembly #3 (bp)	1,533	1,096	520	466
Number of bases aligned to assembly #3	73,659,534	87,869,836	45,221,495	26,182,733
% of unassembled reads aligned to assembly #1 and 2	8.78	1.82	0.29	0.37
Metagenome coverage	0.01	0.01	0.02	0.01

Note: Assemblies were composed of contigs that are >= 300 bp. Contig N50 is defined as the contig length such that using equal or longer contigs produces 50% the bases of the assembly.
Metagenome coverage is calculated as a / (b*c*d) where a = the number of bases in a quality-filtered metagenome, b = 3 Mb average bacterial genome size, c = 400 vaginal bacteria, d = 30 X sequencing depth.

Table 7. The relative abundances of KEGG orthologous groups that were higher in the vaginal microbiomes of both Chinese BV patients compared to non-BV patients (%).

KEGG Orthology ID	Description	BV_SMUB9	BV_SMUB27	NBV_SMUN23	NBV_SMUN25
K00033	6-phosphogluconate dehydrogenase [EC:1.1.1.44]	0.13	0.15	0.05	0.08
K00048	lactaldehyde reductase [EC:1.1.1.77]	0.11	0.12	0	0.09
K00075	UDP-N-acetylmuramate dehydrogenase [EC:1.1.1.158]	0.35	0.36	0	0.28
K00131	glyceraldehyde-3-phosphate dehydrogenase (NADP) [EC:1.2.1.9]	0.15	0.25	0.08	0
K00133	aspartate-semialdehyde dehydrogenase [EC:1.2.1.11]	0.26	0.36	0.25	0.11
K00134	glyceraldehyde 3-phosphate dehydrogenase [EC:1.2.1.12]	0.09	0.05	0.02	0.03
K00239	succinate dehydrogenase flavoprotein subunit [EC:1.3.99.1]	0.54	0.18	0.04	0.05
K00297	methylenetetrahydrofolate reductase (NADPH) [EC:1.5.1.20]	0.21	0.28	0.15	0.12
K00384	thioredoxin reductase (NADPH) [EC:1.8.1.9]	0.36	0.18	0.02	0.11
K00527	ribonucleoside-triphosphate reductase [EC:1.17.4.2]	0.09	0.08	0	0.05
K00528	ferredoxin--NADP+ reductase [EC:1.18.1.2]	0.09	0.13	0	0.01
K00549	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase [EC:2.1.1.14]	0.1	0.11	0.06	0
K00554	tRNA (guanine-N1-)-methyltransferase [EC:2.1.1.31]	0.3	0.25	0.02	0.04
K00604	methionyl-tRNA formyltransferase [EC:2.1.2.9]	0.1	0.14	0	0.02
K00616	transaldolase [EC:2.2.1.2]	0.1	0.15	0.05	0.08
K00625	phosphate acetyltransferase [EC:2.3.1.8]	0.09	0.15	0.03	0.03
K00655	1-acyl-sn-glycerol-3-phosphate acyltransferase [EC:2.3.1.51]	0.21	0.32	0.08	0.17
K00656	formate C-acetyltransferase [EC:2.3.1.54]	0.09	0.16	0.04	0.03
K00760	hypoxanthine phosphoribosyltransferase [EC:2.4.2.8]	0.1	0.1	0	0.09
K00761	uracil phosphoribosyltransferase [EC:2.4.2.9]	0.15	0.07	0	0.03
K00783	hypothetical protein	0.01	0.03	0	0
K00791	tRNA dimethylallyltransferase [EC:2.5.1.75]	0.11	0.13	0.03	0.06
K00796	dihydropteroate synthase [EC:2.5.1.15]	0.16	0.06	0	0.05
K00800	3-phosphoshikimate 1-carboxyvinyltransferase [EC:2.5.1.19]	0.2	0.86	0.06	0.17
K00858	NAD+ kinase [EC:2.7.1.23]	0.13	0.15	0	0.08
K00869	mevalonate kinase [EC:2.7.1.36]	0.21	0.06	0	0.02
K00876	uridine kinase [EC:2.7.1.48]	0.04	0.06	0.02	0.03
K00925	acetate kinase [EC:2.7.2.1]	0.19	0.8	0.08	0.15
K00927	phosphoglycerate kinase [EC:2.7.2.3]	0.14	0.24	0.05	0.13
K00928	aspartate kinase [EC:2.7.2.4]	0.18	0.28	0.02	0.13
K00931	glutamate 5-kinase [EC:2.7.2.11]	0.15	0.02	0	0.01
K00941	hydroxymethylpyrimidine/phosphomethylpyrimidine kinase [EC:2.7.1.49 2.7.4.7]	0.36	0.3	0.03	0.25
K00943	dTMP kinase [EC:2.7.4.9]	0.12	0.11	0.07	0.05
K00949	thiamine pyrophosphokinase [EC:2.7.6.2]	0.16	0.08	0.05	0.07
K00969	nicotinate-nucleotide adenylyltransferase [EC:2.7.7.18]	0.16	0.42	0.1	0.05
K00982	glutamate-ammonia-ligase adenylyltransferase [EC:2.7.7.42]	0.1	0.09	0.02	0.05
K00995	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase [EC:2.7.8.5]	0.15	0.05	0	0.03
K00997	holo-[acyl-carrier protein] synthase [EC:2.7.8.7]	0.06	0.14	0	0.03
K01079	phosphoserine phosphatase [EC:3.1.3.3]	0.12	0.08	0.06	0.02
K01090	protein phosphatase [EC:3.1.3.16]	0.44	0.12	0.04	0.08
K01091	phosphoglycolate phosphatase [EC:3.1.3.18]	0.24	0.45	0.01	0.13
K01092	myo-inositol-1(or 4)-monophosphatase [EC:3.1.3.25]	0.22	0.17	0.06	0.15
K01104	protein-tyrosine phosphatase [EC:3.1.3.48]	0.18	0.24	0.08	0.16
K01139	guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase [EC:3.1.7.2]	0.03	0.02	0	0
K01191	alpha-mannosidase [EC:3.2.1.24]	0.05	0.17	0	0.03
K01239	purine nucleosidase [EC:3.2.2.1]	0.31	0.38	0.19	0.1
K01262	Xaa-Pro aminopeptidase [EC:3.4.11.9]	0.25	0.15	0.03	0.08
K01304	pyroglutamyl-peptidase [EC:3.4.19.3]	0.1	0.62	0.01	0.06
K01338	ATP-dependent Lon protease [EC:3.4.21.53]	0.06	0.04	0.02	0
K01356	repressor LexA [EC:3.4.21.88]	0.14	0.41	0.13	0.09
K01430	urease subunit gamma [EC:3.5.1.5]	0.02	0.01	0	0
K01462	peptide deformylase [EC:3.5.1.88]	0.09	0.08	0.03	0.03
K01516	nucleoside-triphosphatase [EC:3.6.1.15]	0.15	0.05	0.02	0.03
K01533	Cu2+-exporting ATPase [EC:3.6.3.4]	0.05	0.06	0.03	0
K01588	5-(carboxyamino)imidazole ribonucleotide mutase [EC:5.4.99.18]	0.1	0.68	0.04	0.05
K01595	phosphoenolpyruvate carboxylase [EC:4.1.1.31]	0.06	0.05	0.01	0.01
K01597	diphosphomevalonate decarboxylase [EC:4.1.1.33]	0.29	0.07	0.03	0.04
K01676	fumarate hydratase, class I [EC:4.2.1.2]	0.02	0.02	0	0
K01696	tryptophan synthase beta chain [EC:4.2.1.20]	0.17	0.11	0	0.03
K01710	dTDP-glucose 4,6-dehydratase [EC:4.2.1.46]	0.19	0.19	0	0.16
K01714	dihydrodipicolinate synthase [EC:4.2.1.52]	0.15	0.14	0.03	0.08
K01783	ribulose-phosphate 3-epimerase [EC:5.1.3.1]	0.17	0.26	0.1	0.13
K01784	UDP-glucose 4-epimerase [EC:5.1.3.2]	0.16	0.18	0.05	0

Table 7. (cont.)

K01786	L-ribulose-5-phosphate 4-epimerase [EC:5.1.3.4]	0.16	0.01	0	0
K01810	glucose-6-phosphate isomerase [EC:5.3.1.9]	0.13	0.1	0.07	0.05
K01823	isopentenyl-diphosphate delta-isomerase [EC:5.3.3.2]	0.33	0.14	0	0.07
K01834	phosphoglycerate mutase [EC:5.4.2.1]	0.14	0.07	0.05	0.05
K01854	UDP-galactopyranose mutase [EC:5.4.99.9]	0.13	0.2	0.01	0.03
K01858	myo-inositol-1-phosphate synthase [EC:5.5.1.4]	0.1	0.04	0.02	0.03
K01872	alanyl-tRNA synthetase [EC:6.1.1.7]	0.13	0.08	0.05	0.07
K01875	seryl-tRNA synthetase [EC:6.1.1.11]	0.09	0.11	0.03	0
K01880	glycyl-tRNA synthetase [EC:6.1.1.14]	0.11	0.24	0.03	0.08
K01892	histidyl-tRNA synthetase [EC:6.1.1.21]	0.13	0.25	0.05	0.08
K01893	asparaginyl-tRNA synthetase [EC:6.1.1.22]	0.09	0.17	0.03	0
K01915	glutamine synthetase [EC:6.3.1.2]	0.21	0.26	0.06	0.08
K01916	NAD+ synthase [EC:6.3.1.5]	0.02	0.04	0	0
K01929	UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-D-alanine ligase	0.44	0.33	0.01	0.09
K01938	formate--tetrahydrofolate ligase [EC:6.3.4.3]	0.37	0.39	0.03	0.19
K01950	NAD+ synthase (glutamine-hydrolysing) [EC:6.3.5.1]	0.1	0.11	0.03	0.08
K01951	GMP synthase (glutamine-hydrolysing) [EC:6.3.5.2]	0.22	0.3	0.15	0.08
K01956	carbamoyl-phosphate synthase small subunit [EC:6.3.5.5]	0.12	0.13	0.06	0.05
K02006	cobalt/nickel transport system ATP-binding protein	0.22	0.59	0.01	0.12
K02026	multiple sugar transport system permease protein	0.22	0.21	0	0.1
K02027	multiple sugar transport system substrate-binding protein	0.52	0.98	0.01	0.34
K02031	peptide/nickel transport system ATP-binding protein	0.31	0.28	0.24	0.03
K02032	peptide/nickel transport system ATP-binding protein	0.07	0.16	0.06	0.01
K02033	peptide/nickel transport system permease protein	0.25	0.51	0.05	0.09
K02034	peptide/nickel transport system permease protein	0.21	0.18	0.02	0.03
K02035	peptide/nickel transport system substrate-binding protein	0.22	0.21	0.04	0.07
K02039	phosphate transport system protein	0.2	0.3	0.03	0.05
K02115	F-type H+-transporting ATPase subunit gamma [EC:3.6.3.14]	0.1	0.07	0.03	0.03
K02217	ferritin [EC:1.16.3.1]	0.03	0.05	0.02	0
K02283	pilus assembly protein CpaF	0.3	0.15	0	0.13
K02314	replicative DNA helicase [EC:3.6.4.12]	0.28	0.19	0.08	0.08
K02341	DNA polymerase III subunit delta' [EC:2.7.7.7]	0.26	0.18	0.02	0.16
K02342	DNA polymerase III subunit epsilon [EC:2.7.7.7]	0.12	0.04	0	0.03
K02343	DNA polymerase III subunit gamma/tau [EC:2.7.7.7]	0.57	0.16	0.12	0
K02358	elongation factor Tu	0.13	0.11	0.07	0.03
K02483	two-component system, OmpR family, response regulator	0.1	0.14	0	0.08
K02529	LacI family transcriptional regulator	0.16	0.13	0.12	0.05
K02600	N utilization substance protein A	0.06	0.07	0.02	0.03
K02622	topoisomerase IV subunit B [EC:5.99.1.-]	0.05	0.04	0.02	0
K02654	leader peptidase (prepilin peptidase) / N-methyltransferase [EC:3.4.23.43 2.1.1.-]	0.13	0.1	0	0.05
K02838	ribosome recycling factor	0.09	0.1	0.02	0
K02884	large subunit ribosomal protein L19	0.12	0.1	0.05	0.03
K02897	large subunit ribosomal protein L25	0.07	0.03	0.01	0
K02945	small subunit ribosomal protein S1	0.09	0.2	0.05	0
K02948	small subunit ribosomal protein S11	0.07	0.03	0.02	0.01
K03086	RNA polymerase primary sigma factor	0.13	0.36	0.12	0.05
K03088	RNA polymerase sigma-70 factor, ECF subfamily	0.1	0.05	0.03	0
K03148	adenylyltransferase [EC:2.7.7.-]	0.06	0.01	0	0
K03149	thiamine biosynthesis ThiG	0.12	0.02	0	0
K03151	thiamine biosynthesis protein ThiI	0.04	0.04	0	0
K03177	tRNA pseudouridine synthase B [EC:5.4.99.12]	0.22	0.22	0.03	0.02
K03188	urease accessory protein	0.04	0.03	0	0
K03190	urease accessory protein	0.06	0.03	0.02	0
K03293	amino acid transporter, AAT family	0.63	0.56	0.37	0.37
K03299	gluconate:H+ symporter, GntP family	0.49	1.41	0	0.48
K03311	branched-chain amino acid:cation transporter, LIVCS family	0.19	0.21	0.16	0.11
K03320	ammonium transporter, Amt family	0.3	0.17	0.03	0.11
K03424	TatD DNase family protein [EC:3.1.21.-]	0.15	0.07	0.03	0.05
K03469	ribonuclease HI [EC:3.1.26.4]	0.12	0.13	0.04	0.01
K03470	ribonuclease HII [EC:3.1.26.4]	0.27	0.09	0.05	0.06
K03496	chromosome partitioning protein	0.33	0.42	0.18	0.31
K03497	chromosome partitioning protein, ParB family	0.08	0.04	0.02	0
K03501	ribosomal RNA small subunit methyltransferase G [EC:2.1.1.170]	0.2	0.3	0	0.12
K03525	type III pantothenate kinase [EC:2.7.1.33]	0.2	0.37	0.03	0.16
K03531	cell division protein FtsZ	0.17	0.29	0.13	0.11
K03536	ribonuclease P protein component [EC:3.1.26.5]	0.05	0.03	0	0
K03545	trigger factor	0.15	0.21	0.03	0.04
K03550	holliday junction DNA helicase RuvA	0.12	0.11	0.09	0.01

Table 7. (cont.)

K03551	holliday junction DNA helicase RuvB	0.19	0.13	0.02	0.02
K03575	A/G-specific adenine glycosylase [EC:3.2.2.-]	0.13	0.28	0	0.03
K03578	ATP-dependent helicase HrpA [EC:3.6.4.13]	0.2	0.22	0.07	0.13
K03589	cell division protein FtsQ	0.09	0.09	0	0.03
K03621	glycerol-3-phosphate acyltransferase PlsX [EC:2.3.1.15]	0.04	0.05	0.02	0
K03624	transcription elongation factor GreA	0.14	0.24	0.01	0.04
K03629	DNA replication and repair protein RecF	0.14	0.11	0.04	0.06
K03631	DNA repair protein RecN (Recombination protein N)	0.24	0.05	0.03	0.01
K03648	uracil-DNA glycosylase [EC:3.2.2.-]	0.1	0.11	0.02	0.04
K03655	ATP-dependent DNA helicase RecG [EC:3.6.4.12]	0.3	0.13	0	0.03
K03684	ribonuclease D [EC:3.1.13.5]	0.1	0.14	0.02	0.08
K03686	molecular chaperone DnaJ	0.19	0.05	0.02	0
K03696	ATP-dependent Clp protease ATP-binding subunit ClpC	0.69	0.43	0.32	0.34
K03699	putative hemolysin	0.13	0.25	0.11	0.03
K03701	excinuclease ABC subunit A	0.03	0.01	0	0
K03703	excinuclease ABC subunit C	0.22	0.05	0.02	0
K03704	cold shock protein (beta-ribbon, CspA family)	0.06	0.14	0.02	0.04
K03707	thiaminase (transcriptional activator TenA) [EC:3.5.99.2]	0.14	0.29	0	0
K03734	thiamine biosynthesis lipoprotein	0.25	0.13	0	0.06
K03742	competence/damage-inducible protein CinA	0.04	0.08	0.02	0
K03743	N/A	0.13	0.04	0	0
K03744	LemA protein	0.13	0.2	0	0.07
K03786	3-dehydroquinate dehydratase II [EC:4.2.1.10]	0.09	0.09	0.01	0.04
K03789	ribosomal-protein-alanine N-acetyltransferase [EC:2.3.1.128]	0.12	0.21	0.03	0.08
K03798	cell division protease FtsH [EC:3.4.24.-]	0.25	0.12	0	0.07
K03799	heat shock protein HtpX [EC:3.4.24.-]	0.3	0.36	0	0.06
K03977	GTP-binding protein	0.07	0.12	0.04	0.05
K03980	putative peptidoglycan lipid II flippase	0.4	0.17	0.15	0
K04042	bifunctional UDP-N-acetylglucosamine pyrophosphorylase / Glucosamine-1-phosphate N-ac	0.11	0.09	0.02	0.05
K04066	primosomal protein N' (replication factor Y) (superfamily II helicase) [EC:3.6.4.-]	0.15	0.26	0.03	0.03
K04069	pyruvate formate lyase activating enzyme [EC:1.97.1.4]	0.15	0.15	0.05	0.08
K04075	tRNA(Ile)-lysidine synthase [EC:6.3.4.-]	0.24	0.17	0.03	0.05
K04092	chorismate mutase [EC:5.4.99.5]	0.08	0.04	0	0
K04095	cell filamentation protein	0.19	0.05	0	0.04
K04096	DNA processing protein	0.35	0.17	0.02	0.03
K04485	DNA repair protein RadA/Sms	0.34	0.62	0.12	0.12
K04517	prephenate dehydrogenase [EC:1.3.1.12]	0.17	0.1	0	0.08
K05364	peptidoglycan glycosyltransferase [EC:2.4.1.129]	0.4	0.59	0.36	0.19
K06024	segregation and condensation protein B	0.01	0.02	0	0
K06148	ATP-binding cassette, subfamily C, bacterial	0.04	0.06	0.02	0
K06180	ribosomal large subunit pseudouridine synthase D [EC:5.4.99.12]	0.12	0.12	0.03	0
K06187	recombination protein RecR	0.16	0.17	0.02	0.05
K06199	CrcB protein	0.21	0.11	0	0.08
K06213	magnesium transporter	0.06	0.06	0.03	0
K06215	pyridoxine biosynthesis protein [EC:4.-.-.-]	0.27	0.55	0.15	0.18
K06217	phosphate starvation-inducible protein PhoH and related proteins	0.18	0.14	0	0.08
K06881	N/A	0.05	0.05	0.02	0
K06942	N/A	0.08	0.07	0.03	0.03
K06950	uncharacterized protein	0.21	0.05	0.02	0.02
K06966	N/A	0.21	0.22	0.03	0.14
K07012	N/A	0.44	1.07	0.01	0.06
K07021	N/A	0.07	0.08	0.03	0
K07024	N/A	0.16	0.15	0.03	0.07
K07043	N/A	0.02	0.05	0.01	0
K07056	N/A	0.26	0.14	0.03	0
K07089	N/A	0.13	0.64	0	0.03
K07114	uncharacterized protein	0.27	0.04	0.01	0.02
K07133	N/A	0.14	0.3	0.08	0.01
K07183	response regulator NasT	0.24	0.32	0.13	0.07
K07230	N/A	0.16	0.19	0.08	0.07
K07243	high-affinity iron transporter	0.23	0.07	0	0.06
K07259	D-alanyl-D-alanine carboxypeptidase / D-alanyl-D-alanine-endopeptidase (penicillin-binding	0.41	0.12	0.01	0.04
K07478	putative ATPase	0.27	0.59	0.15	0.09
K07493	putative transposase	0.05	0.05	0.02	0
K07586	hypothetical protein	0.03	0.04	0.01	0
K07862	serine/threonine transporter	0.25	0.21	0	0.1
K08289	phosphoribosylglycinamide formyltransferase 2 [EC:2.1.2.2]	0.07	0.07	0.02	0.05
K08372	putative serine protease PepD [EC:3.4.21.-]	0.08	0.07	0.04	0.04

Table 7. (cont.)

K08591	glycerol-3-phosphate acyltransferase PlsY [EC:2.3.1.15]	0.05	0.04	0.02	0
K08659	dipeptidase [EC:3.4.-.-]	0.23	0.12	0.08	0.07
K08681	glutamine amidotransferase [EC:2.6.-.-]	0.14	0.28	0.07	0.08
K08884	serine/threonine protein kinase, bacterial [EC:2.7.11.1]	0.35	0.15	0.03	0
K09710	ribosome-associated protein	0.1	0.05	0	0
K09761	ribosomal RNA small subunit methyltransferase E [EC:2.1.1.-]	0.1	0.11	0.03	0.09
K09772	cell division inhibitor SepF	0.1	0.16	0.03	0.05
K09903	uridylate kinase [EC:2.7.4.22]	0.09	0.08	0.01	0
K10005	glutamate transport system substrate-binding protein	0.19	0.26	0.06	0.06
K10006	glutamate transport system permease protein	0.15	0.23	0.08	0.04
K10007	glutamate transport system permease protein	0.24	0.26	0.05	0.15
K10008	glutamate transport system ATP-binding protein [EC:3.6.3.-]	0.08	0.12	0.03	0.05
K10439	ribose transport system substrate-binding protein	0.18	0.08	0	0.05
K10440	ribose transport system permease protein	0.23	0.06	0	0.03
K10975	allantoin permease	0.12	0.04	0.02	0
K11031	thiol-activated cytolysin	0.1	0.09	0.06	0.01
K11753	riboflavin kinase / FMN adenylyltransferase [EC:2.7.1.26 2.7.7.2]	0.3	0.37	0.24	0
K11754	dihydrofolate synthase / folylpolyglutamate synthase [EC:6.3.2.12 6.3.2.17]	0.32	0.24	0.19	0.05
K12257	SecD/SecF fusion protein	0.1	0.15	0.09	0
K13640	MerR family transcriptional regulator, heat shock protein HspR	0.16	0.04	0	0
K13787	geranylgeranyl diphosphate synthase, type I [EC:2.5.1.1 2.5.1.10 2.5.1.29]	0.11	0.21	0.04	0.09
K13788	phosphate acetyltransferase [EC:2.3.1.8]	0.12	0.22	0.03	0.07
K17330	N/A	0.15	0.01	0	0
K18682	N/A	0.01	0.18	0	0
Note: N/A represents there is no related information.					

Table 8. Taxonomic abundance in Chinese women with and without BV at the genus level (%).

Bacterial genera	Chinese women with BV	Chinese women without BV
Lactobacillus	55.9 ± 5.49	74.16 ± 10.92
Acinetobacter	0.43 ± 0.09	0.4 ± 0.35
Adhaeribacter	0.04 ± 0.03	0 ± 0.00
Aerococcus	5.68 ± 1.87	2.86 ± 2.84
Aeromonas	0.35 ± 0.20	0.83 ± 0.83
Alkanindiges	0.01 ± 0.00	0 ± 0.00
Anaerococcus	0.01 ± 0.01	0.01 ± 0.00
Atopobium	0.12 ± 0.03	0.22 ± 0.21
Bacillus	0.03 ± 0.03	0 ± 0.00
Bacteroides	0.05 ± 0.03	0 ± 0.00
Bifidobacterium	0.01 ± 0.00	0 ± 0.00
Brevundimonas	0.02 ± 0.00	0 ± 0.00
Campylobacter	0 ± 0.00	0.04 ± 0.04
Chryseobacterium	0.01 ± 0.01	0 ± 0.00
Cloacibacterium	0.37 ± 0.15	0.01 ± 0.00
Comamonas	0.07 ± 0.04	0.07 ± 0.07
Corynebacterium	0.02 ± 0.01	0.02 ± 0.02
Deinococcus	0.08 ± 0.04	0 ± 0.00
Dialister	5.57 ± 1.86	1 ± 0.50
Diaphorobacter	0.44 ± 0.19	0.02 ± 0.01
Enterococcus	2.44 ± 0.90	0.14 ± 0.07
Gardnerella	1.37 ± 0.28	1.01 ± 0.76
Gemella	0.01 ± 0.01	0.01 ± 0.01
Gp4	0.06 ± 0.03	0 ± 0.00
Hallella	0.06 ± 0.04	0.06 ± 0.06
Halomonas	0.16 ± 0.16	0 ± 0.00
Hyphomicrobium	0.01 ± 0.00	0 ± 0.00
Janthinobacterium	0.02 ± 0.02	0 ± 0.00
Jonquetella	0.18 ± 0.14	0 ± 0.00
Mycoplasma	0.01 ± 0.00	0.01 ± 0.00
Neisseria	0 ± 0.00	0.01 ± 0.00
OD1 genera incertae sedis	0.01 ± 0.01	0.03 ± 0.03
Parvimonas	0.01 ± 0.00	0 ± 0.00
Peptoniphilus	0.05 ± 0.02	0 ± 0.00
Peptostreptococcus	0.03 ± 0.02	0.09 ± 0.06
Prevotella	10.78 ± 3.04	1.31 ± 1.00
Propionibacterium	0.03 ± 0.01	0 ± 0.00
Pseudomonas	0 ± 0.00	7.83 ± 7.78
Saccharofermentans	0.92 ± 0.61	0.01 ± 0.00
Sneathia	2.15 ± 1.15	2.78 ± 2.74
Staphylococcus	2.08 ± 0.95	1.36 ± 1.32
Streptococcus	2.03 ± 1.70	3.63 ± 3.63
Ureaplasma	8.28 ± 2.93	2.08 ± 1.97
Veillonella	0.1 ± 0.06	0.01 ± 0.00
Xanthobacter	0.01 ± 0.01	0 ± 0.00
Subdivision3 genera incertae sedis	0 ± 0.00	0.01 ± 0.01

Note: Numbers indicate the number of reads assigned to that bacterial genus (mean ± standard error of the mean).

Table 9. Taxonomic abundance in Chinese women with and without BV at the strain level (%).

Bacterial strains	Chinese women with BV	Chinese women without BV
Anaerococcus hydrogenalis ACS-025-V-Sch4	0.04 ± 0.04	0 ± 0.00
Anaerococcus lactolyticus ATCC 51172	0 ± 0.00	0.01 ± 0.01
Atopobium vaginae DSM 15829	4.33 ± 4.27	0.07 ± 0.04
Atopobium vaginae PB189-T1-4	0.04 ± 0.01	0.01 ± 0.01
Bifidobacterium dentium JCVIHMP022	0.01 ± 0.00	41.43 ± 9.84
Clostridiales genomosp. BVAB3 str. UPII9-5	0.11 ± 0.01	0.04 ± 0.02
Corynebacterium glucuronolyticum ATCC 51867	0.01 ± 0.01	0.74 ± 0.71
Corynebacterium lipophiloflavum DSM 44291	0 ± 0.00	1.3 ± 0.85
Corynebacterium pseudogenitalium ATCC 33035	0.01 ± 0.01	0 ± 0.00
Dialister microaerophilus UPII 345-E	0.02 ± 0.01	0.03 ± 0.02
Enterococcus faecalis ERV85	0.06 ± 0.06	0 ± 0.00
Eremococcus coleocola ACS-139-V-Col8	0 ± 0.00	0.01 ± 0.01
Escherichia coli 83972	0 ± 0.00	7.58 ± 6.27
Finegoldia magna BVS033A4	0.01 ± 0.01	0 ± 0.00
Finegoldia magna SY403409CC001050417	0.01 ± 0.01	0 ± 0.00
Gardnerella vaginalis 315-A	16.9 ± 0.84	1.7 ± 1.33
Gardnerella vaginalis 409-05	28.12 ± 2.72	4.52 ± 1.85
Gardnerella vaginalis ATCC 14019	17.02 ± 1.79	1.61 ± 1.13
Gardnerella vaginalis HMP9231	30.88 ± 3.28	3.13 ± 2.19
Klebsiella oxytoca 10-5245	0 ± 0.00	0.01 ± 0.01
Lactobacillus gasseri 224-1	0.01 ± 0.00	29.17 ± 4.62
Lactobacillus iners ATCC 55195	0.02 ± 0.01	0.17 ± 0.16
Lactobacillus iners DSM 13335	0.01 ± 0.00	0.14 ± 0.14
Lactobacillus iners LactinV 01V1-a	0.02 ± 0.01	0.19 ± 0.19
Lactobacillus iners LactinV 03V1-b	0.05 ± 0.00	0.33 ± 0.33
Lactobacillus iners LactinV 09V1-c	0.07 ± 0.03	0.29 ± 0.29
Lactobacillus iners LactinV 11V1-d	0.02 ± 0.01	0.29 ± 0.29
Lactobacillus iners LEAF 2052A-d	0.01 ± 0.01	0.09 ± 0.09
Lactobacillus iners LEAF 2053A-b	0.01 ± 0.00	0.14 ± 0.14
Lactobacillus iners LEAF 2062A-h1	0.08 ± 0.02	0.14 ± 0.14
Lactobacillus iners LEAF 3008A-a	0.02 ± 0.01	0.31 ± 0.31
Lactobacillus iners SPIN 1401G	0.04 ± 0.01	0.51 ± 0.51
Lactobacillus iners SPIN 2503V10-D	0.05 ± 0.02	0.55 ± 0.55
Lactobacillus iners UPII 143-D	0.04 ± 0.01	0.66 ± 0.66
Lactobacillus iners UPII 60-B	0.13 ± 0.03	1.27 ± 1.27
Lactobacillus jensenii SJ-7A-US	0.08 ± 0.07	0 ± 0.00
Lactobacillus salivarius ACS-116-V-Col5a	0.01 ± 0.00	0 ± 0.00
Lactobacillus vaginalis ATCC 49540	0 ± 0.00	3.46 ± 1.35
Megasphaera genomosp type 1 str 28L	0.04 ± 0.00	0.01 ± 0.01
Megasphaera sp. UPII 135-E	0 ± 0.00	0.01 ± 0.01
Mobiluncus curtisii ATCC 43063	0.03 ± 0.03	0.01 ± 0.01
Mobiluncus curtisii ATCC 51333	0.07 ± 0.07	0.03 ± 0.00
Mobiluncus curtisii subsp curtisii ATCC 35241	0.01 ± 0.01	0.01 ± 0.01
Mobiluncus curtisii subsp holmesii ATCC 35242	0.01 ± 0.01	0.01 ± 0.01
Mobiluncus mulieris 28-1	0.02 ± 0.02	0.01 ± 0.00

Table 9. (cont.)

Mobiluncus mulieris ATCC 35239	0.01 ± 0.01	0 ± 0.00
Mobiluncus mulieris FB024-16	0.01 ± 0.01	0 ± 0.00
Peptoniphilus duerdenii ATCC BAA-1640	0.01 ± 0.01	0.01 ± 0.01
Peptoniphilus harei ACS-146-V-Sch2b	0.01 ± 0.01	0.02 ± 0.02
Peptoniphilus lacrimalis 315-B	0.04 ± 0.04	0 ± 0.00
Prevotella amnii CRIS 21A-A	0.01 ± 0.01	0.01 ± 0.01
Prevotella bivia JCVIHMP010	0.05 ± 0.01	0 ± 0.00
Prevotella buccalis ATCC 35310	0.06 ± 0.04	0 ± 0.00
Prevotella denticola CRIS 18C-A	0.02 ± 0.01	0 ± 0.00
Prevotella disiens FB035-09AN	0.06 ± 0.02	0 ± 0.00
Prevotella oralis ATCC 33269	0.01 ± 0.01	0 ± 0.00
Prevotella timonensis CRIS 5C-B1	1.27 ± 0.92	0.01 ± 0.01
Staphylococcus epidermidis BVS058A4	0.08 ± 0.08	0 ± 0.00
Streptococcus bovis ATCC 700338	0.01 ± 0.01	0.04 ± 0.04
Streptococcus pseudoporcinus SPIN 20026	0.04 ± 0.01	0.04 ± 0.03
Streptococcus urinalis FB127-CNA-2	0.01 ± 0.01	0 ± 0.00

Note: Numbers indicate the number of reads assigned to that bacterial genus (mean ± standard error of the mean).

Table 10. Sample descriptions.

Sample ID	403	405	409	412	416
Disease Status	NBV	NBV	Asymptomatic BV	Symptomatic BV	Symptomatic BV
BV Symptoms	NA	NA	NA	DPI	D
Amsel Criteria	Amsel-negative	Amsel-negative	Amsel-negative	Amsel-positive	Amsel-positive
Nugent Score	6	0	9	6	3
Race/Ethnicity	Asian American/Indian	Caucasian	Caucasian	African American	Caucasian
Age In Years	20	36-40	22	26	22
Length Of Residence	2-5 years	More than 10 years	2-5 years	5-10 years	More than 10 years
Marital Status	Single	Divorced	Single	Single	Single
Sexual Orientation	Bisexual	Heterosexual	Heterosexual	Not reported	Heterosexual
Days Since Last Period	16	Not reported	30	15	20
Tampon Use During Periods	Yes	Yes	Yes	No	Yes
Number Of Previous Pregnancies	0	0	0	2	0
Number Of Sexual Partners In the Past 6 Months	2	0	4-6	1	1
Gender of Sexual Partners In the Past 6 Months	Male	Not reported	Male	Male	Male
New Sexual Partner In the Past 6 Months	Yes	No	Yes	No	No
Days Since Intercourse	0	Not reported	3	10	3
Number Of Vaginal Sex/Week	7+	0	0-2	7+	3-4
Number Of Oral Sex/Week	5-6	0	0-2	0	0-2
Number Of Anal Sex/Week	3-4	0	0-2	0	0
Condom Use	More than half of the time	Not reported	Less than half of the time	More than half of the time	No
Number Of Bath/Week	0-2	0-2	0-2	5-6	0-2
Days Since Last Douche	90			7	6
Comorbidities				Asthma	

Note: BV in the USA was determined by the Amsel or Nugent criteria. Amsel-positive patients are named as symptomatic BV patients. Amsel-negative patients with high Nugent Score (7-10) are named as asymptomatic BV patients. Symptoms: (D)ischarge, (O)dor, (P)ain, (I)tching. NA represents that there are no BV symptoms observed in the individual.

Table 11. 16S rRNA sequence statistics of vaginal samples.

Sample ID	Total sequences number	Quality-filtered sequences number	Average quality-filtered sequence length (nt)
403	99,431	63,459	440
405	95,599	79,011	488
409	92,029	54,495	475
412	232,634	158,473	377
416	105,012	81,738	493

Table 12. Summary of the percentage of bases covered for each HMP urogenital tract reference genome in each sample.

	Symptomatic_ BV_USA_412	Symptomatic_ BV_USA_416	Asymptomatic _BV_USA_409	NBV_USA _403	NBV_USA_405
Acinetobacter_baumannii_ATCC_19606_=_CIP_70.34	0.01%	0.00%	0.05%	0.15%	0.04%
Actinobaculum_massiliae_ACS-171-V-Col2	0.00%	0.00%	0.00%	0.00%	0.00%
Actinomyces_coleocanis_DSM_15436	0.00%	0.00%	0.00%	0.00%	0.00%
Actinomyces_neuii_BVS029A5	0.00%	0.02%	0.00%	0.32%	0.00%
Actinomyces_turicensis_ACS-279-V-Col4	0.00%	0.00%	0.04%	0.00%	0.00%
Actinomyces_urogenitalis_DSM_15434	0.00%	0.00%	0.00%	0.00%	0.00%
Aerococcus_urinae_ACS-120-V-Col10a	0.01%	0.00%	0.07%	0.03%	0.00%
Aerococcus_viridans_ATCC_11563	0.00%	0.00%	0.02%	0.00%	0.00%
Anaerococcus_hydrogenalis_ACS-025-V-Sch4	0.01%	0.03%	0.07%	17.86%	0.02%
Anaerococcus_lactolyticus_ATCC_51172	0.01%	0.01%	0.25%	0.41%	0.01%
Anaerococcus_prevotii_ACS-065-V-Col13	0.00%	0.00%	0.05%	0.17%	0.02%
Anaerococcus_tetradius_ATCC_35098	0.01%	0.03%	0.15%	0.07%	0.00%
Atopobium_vaginae_DSM_15829	0.36%	0.15%	92.09%	86.39%	0.47%
Atopobium_vaginae_PB189-T1-4	0.00%	0.00%	1.41%	0.23%	0.01%
Bifidobacterium_breve_ACS-071-V-Sch8b	0.00%	0.00%	0.00%	0.00%	0.00%
Bifidobacterium_dentium_ATCC_27679	0.00%	0.00%	0.00%	0.00%	0.00%
Bifidobacterium_dentium_JCVIHMP022	0.00%	0.00%	0.00%	0.00%	0.00%
Brevibacterium_mcbrellneri_ATCC_49030	0.00%	0.00%	0.00%	0.00%	0.00%
Chryseobacterium_gleum_ATCC_35910	0.00%	0.00%	0.00%	0.00%	0.00%
Clostridiales_genomosp._BVAB3_str._UPII9-5	0.17%	0.01%	1.04%	0.65%	0.00%
Corynebacterium_aurimucosum_ATCC_700975	0.00%	0.02%	0.00%	0.01%	0.00%
Corynebacterium_genitalium_ATCC_33030	0.00%	0.00%	0.00%	0.00%	0.00%
Corynebacterium_glucuronolyticum_ATCC_51866	0.00%	0.00%	0.00%	0.01%	0.00%
Corynebacterium_glucuronolyticum_ATCC_51867	0.00%	0.00%	0.00%	0.00%	0.00%
Corynebacterium_jeikeium_ATCC_43734	0.00%	0.00%	0.00%	0.00%	0.00%
Corynebacterium_lipophiloflavum_DSM_44291	0.00%	0.00%	0.00%	0.00%	0.00%
Corynebacterium_pseudogenitalium_ATCC_33035	0.01%	0.10%	0.01%	0.79%	0.09%
Corynebacterium_striatum_ATCC_6940	0.00%	0.01%	0.00%	0.02%	0.01%
Dialister_micraerophilus_DSM_19965	0.30%	0.15%	7.12%	3.22%	0.04%
Dialister_microaerophilus_UPII_345-E	0.01%	0.01%	10.51%	0.25%	0.00%
Enterococcus_durans_FB129-CNAB-4	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_ATCC_29200	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_ERV81	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_ERV85	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_ERV93	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_HH22	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_TX0312	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_TX0635	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_TX0855	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecium_E422	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecium_P1139	0.00%	0.00%	0.06%	0.02%	0.00%
Enterococcus_faecium_S447	0.00%	0.00%	0.00%	0.00%	0.00%
Eremococcus_coleocola_ACS-139-V-Col8	0.00%	0.00%	0.01%	0.00%	0.00%
Escherichia_coli_83972	0.01%	0.00%	0.00%	0.01%	0.05%
Finegoldia_magna_ACS-171-V-Col3	0.00%	0.00%	0.24%	0.12%	0.03%
Finegoldia_magna_ATCC_53516	0.00%	0.34%	0.18%	0.16%	0.02%
Finegoldia_magna_BVS033A4	0.00%	0.00%	0.35%	0.02%	0.01%
Finegoldia_magna_SY403409CC001050417	0.01%	0.04%	0.28%	1.27%	0.05%
Fusobacterium_nucleatum_subsp._nucleatum_ATCC_23726	0.04%	0.01%	0.12%	0.02%	0.00%
Gardnerella_vaginalis_315-A	0.64%	0.10%	6.83%	19.59%	0.05%
Gardnerella_vaginalis_409-05	0.95%	1.93%	99.24%	87.03%	1.53%
Gardnerella_vaginalis_ATCC_14019	0.42%	0.05%	7.18%	17.61%	0.02%
Gardnerella_vaginalis_HMP9231	0.87%	0.08%	7.11%	17.06%	0.05%

Table 12. (cont.)

Haemophilus_parainfluenzae_HK262	0.00%	0.00%	0.00%	3.37%	0.00%
Klebsiella_oxytoca_10-5245	0.00%	0.00%	0.00%	0.02%	0.00%
Klebsiella_oxytoca_10-5250	0.00%	0.00%	0.00%	0.00%	0.00%
Lactobacillus_coleohominis_101-4-CHN	1.32%	0.11%	0.00%	0.24%	0.35%
Lactobacillus_crispatus_125-2-CHN	0.14%	0.19%	0.00%	0.00%	1.86%
Lactobacillus_crispatus_214-1	0.04%	0.23%	0.01%	0.02%	0.90%
Lactobacillus_crispatus_CTV-05	0.46%	2.36%	0.00%	0.05%	3.23%
Lactobacillus_crispatus_FB049-03	0.67%	0.45%	0.01%	0.00%	1.38%
Lactobacillus_crispatus_FB077-07	0.48%	0.55%	0.00%	0.01%	3.79%
Lactobacillus_crispatus_JV-V01	0.09%	0.02%	0.00%	0.00%	0.08%
Lactobacillus_crispatus_MV-1A-US	0.41%	0.80%	0.00%	0.00%	0.22%
Lactobacillus_crispatus_MV-3A-US	1.22%	0.53%	0.00%	0.00%	0.32%
Lactobacillus_crispatus_SJ-3C-US	0.42%	0.23%	0.00%	0.02%	0.99%
Lactobacillus_delbrueckii_subsp._bulgaricus_PB2003/044-T3-4	0.00%	0.00%	0.00%	0.00%	0.00%
Lactobacillus_fermentum_28-3-CHN	0.00%	0.02%	0.00%	0.01%	0.02%
Lactobacillus_gasseri_202-4	0.74%	0.64%	0.02%	16.71%	0.02%
Lactobacillus_gasseri_224-1	0.05%	0.04%	0.02%	0.17%	0.03%
Lactobacillus_gasseri_JV-V03	1.63%	0.02%	0.00%	0.18%	0.02%
Lactobacillus_gasseri_MV-22	0.02%	0.04%	0.00%	0.07%	0.23%
Lactobacillus_gasseri_SJ-9E-US	0.06%	0.00%	0.00%	0.01%	0.01%
Lactobacillus_gasseri_SV-16A-US	0.02%	0.02%	0.00%	0.85%	0.00%
Lactobacillus_iners_ATCC_55195	2.79%	1.61%	1.42%	0.01%	0.71%
Lactobacillus_iners_DSM_13335	3.31%	2.09%	0.75%	0.01%	0.48%
Lactobacillus_iners_LactinV_01V1-a	1.65%	1.64%	0.86%	0.00%	0.19%
Lactobacillus_iners_LactinV_03V1-b	2.57%	1.70%	0.95%	0.04%	0.26%
Lactobacillus_iners_LactinV_09V1-c	2.74%	1.88%	0.99%	0.00%	0.27%
Lactobacillus_iners_LactinV_11V1-d	2.44%	1.35%	0.80%	0.03%	0.53%
Lactobacillus_iners_LEAF_2052A-d	2.60%	1.41%	1.38%	0.00%	0.68%
Lactobacillus_iners_LEAF_2053A-b	9.22%	8.26%	1.22%	0.23%	0.79%
Lactobacillus_iners_LEAF_2062A-h1	3.18%	1.45%	1.06%	0.04%	0.19%
Lactobacillus_iners_LEAF_3008A-a	1.55%	1.67%	0.53%	0.00%	0.33%
Lactobacillus_iners_SPIN_1401G	1.79%	1.37%	0.72%	0.91%	0.41%
Lactobacillus_iners_SPIN_2503V10-D	1.60%	3.15%	0.60%	0.00%	0.26%
Lactobacillus_iners_UPII_143-D	1.63%	1.15%	0.82%	0.00%	0.20%
Lactobacillus_iners_UPII_60-B	4.02%	3.25%	1.21%	0.13%	0.47%
Lactobacillus_jensenii_1153	0.00%	1.45%	0.00%	0.03%	0.00%
Lactobacillus_jensenii_115-3-CHN	0.00%	0.00%	0.00%	0.00%	0.02%
Lactobacillus_jensenii_269-3	0.00%	0.35%	0.00%	0.01%	0.00%
Lactobacillus_jensenii_27-2-CHN	0.00%	0.00%	0.00%	0.00%	0.02%
Lactobacillus_jensenii_JV-V16	0.12%	0.05%	0.00%	0.00%	9.02%
Lactobacillus_jensenii_SJ-7A-US	0.20%	6.46%	0.01%	0.01%	0.12%
Lactobacillus_johnsonii_ATCC_33200	0.00%	0.00%	0.00%	0.11%	0.00%
Lactobacillus_oris_PB013-T2-3	0.06%	0.02%	0.00%	0.00%	0.01%
Lactobacillus_salivarius_ACS-116-V-Col5a	0.00%	0.00%	0.00%	0.01%	0.00%
Lactobacillus_vaginalis_ATCC_49540	0.02%	1.10%	0.01%	1.51%	2.00%
Megasphaera_genomosp._type_1_str._28L	0.00%	0.00%	0.17%	0.07%	0.00%
Megasphaera_sp._UPII_135-E	0.01%	0.00%	0.46%	0.05%	0.00%
Megasphaera_sp._UPII_199-6	0.00%	0.00%	0.15%	0.04%	0.00%
Mobiluncus_curtisii_ATCC_43063	0.00%	0.00%	0.00%	0.00%	0.00%
Mobiluncus_curtisii_ATCC_51333	0.00%	0.00%	0.00%	0.00%	0.00%
Mobiluncus_curtisii_subsp._curtisii_ATCC_35241	0.00%	0.00%	0.00%	0.00%	0.00%
Mobiluncus_curtisii_subsp._holmesii_ATCC_35242	0.00%	0.00%	0.01%	0.02%	0.00%
Mobiluncus_mulieris_28-1	0.00%	0.00%	0.00%	0.00%	0.00%
Mobiluncus_mulieris_ATCC_35239	0.00%	0.00%	0.00%	0.00%	0.00%
Mobiluncus_mulieris_ATCC_35243	0.00%	0.00%	0.00%	0.00%	0.00%

Table 12. (cont.)

Mobiluncus_mulieris_FB024-16	0.00%	0.00%	0.00%	0.00%	0.00%
Mycobacterium_parascrofulaceum_ATCC_BAA-614	0.00%	0.00%	0.00%	0.00%	0.00%
Myroides_odoratimimus_CCUG_12901	0.00%	0.00%	0.00%	0.00%	0.00%
Myroides_odoratimimus_CCUG_3837	0.00%	0.00%	0.00%	0.00%	0.00%
Pasteurella_bettyae_CCUG_2042	0.00%	0.00%	0.02%	0.02%	0.00%
Peptoniphilus_duerdenii_ATCC_BAA-1640	0.06%	0.00%	0.25%	0.08%	0.00%
Peptoniphilus_harei_ACS-146-V-Sch2b	0.00%	0.11%	0.21%	8.29%	0.11%
Peptoniphilus_lacrimalis_315-B	0.13%	0.00%	0.28%	0.06%	0.01%
Porphyromonas_asaccharolytica_PR426713P-I	0.00%	0.00%	0.00%	0.01%	0.00%
Porphyromonas_uenonis_60-3	0.01%	0.00%	0.67%	0.00%	0.00%
Prevotella_amnii_CRIS_21A-A	0.05%	0.02%	85.27%	0.31%	0.11%
Prevotella_bivia_JCVIHMP010	0.08%	2.14%	84.10%	55.79%	0.81%
Prevotella_buccalis_ATCC_35310	0.05%	0.01%	9.99%	0.15%	0.01%
Prevotella_denticola_CRIS_18C-A	0.00%	0.01%	1.58%	0.20%	0.00%
Prevotella_disiens_FB035-09AN	0.05%	0.02%	1.80%	0.28%	0.01%
Prevotella_oralis_ATCC_33269	0.00%	0.00%	0.07%	0.01%	0.00%
Prevotella_timonensis_CRIS_5C-B1	0.14%	0.02%	1.76%	0.26%	0.15%
Propionibacterium_sp._409-HC1	0.00%	0.02%	0.00%	0.00%	0.00%
Propionibacterium_sp._434-HC2	0.00%	0.00%	0.00%	0.01%	0.00%
Proteus_mirabilis_ATCC_29906	0.00%	0.00%	0.00%	0.01%	0.00%
Roseomonas_cervicalis_ATCC_49957	0.00%	0.00%	0.00%	0.00%	0.00%
Sphingobacterium_spiritivorum_ATCC_33300	0.00%	0.00%	0.00%	0.00%	0.00%
Sphingobacterium_spiritivorum_ATCC_33861	0.00%	0.00%	0.00%	0.00%	0.00%
Staphylococcus_aureus_subsp._aureus_MN8	0.00%	0.00%	0.02%	0.02%	0.05%
Staphylococcus_epidermidis_BVS058A4	0.07%	0.41%	0.01%	0.04%	0.06%
Staphylococcus_lugdunensis_ACS-027-V-Sch2	0.00%	0.00%	0.00%	0.00%	0.00%
Staphylococcus_simulans_ACS-120-V-Sch1	0.00%	0.00%	0.00%	0.00%	0.00%
Streptococcus_bovis_ATCC_700338	0.00%	0.00%	0.02%	0.01%	0.00%
Streptococcus_pseudoporcinus_SPIN_20026	0.01%	0.01%	0.08%	0.03%	0.00%
Streptococcus_urinalis_FB127-CNA-2	0.02%	0.00%	0.07%	0.01%	0.00%
Treponema_phagedenis_F0421	0.00%	0.00%	0.01%	0.00%	0.00%
Veillonella_atypica_ACS-049-V-Sch6	0.00%	0.00%	0.00%	0.02%	0.00%
Veillonella_atypica_ACS-134-V-Col7a	0.00%	0.00%	0.03%	0.08%	0.00%
Veillonella_parvula_ACS-068-V-Sch12	0.00%	0.03%	0.06%	0.26%	0.00%
Veillonella_ratti_ACS-216-V-Col6b	0.02%	0.00%	0.02%	0.02%	0.00%
Note: only reads (mapping quality > 20) were considered as correct alignment.					

Table 13. List of scaffolds whose sequencing depths were bigger than 20X in the US BV patients and which were classified by essential single copy gene.

Taxonomic ID	Scaffold ID	Scaffold length	GC (%)	Sequencing depth in subject 412	Sequencing depth in subject 416
Firmicutes	9	96002	33.84	21.105X	29.683X
Firmicutes	44	15377	34.96	30.507X	24.649X
Firmicutes	229	9947	32.71	31.366X	25.312X
Firmicutes	1162	8425	32.45	28.363X	20.202X
Firmicutes	2886	12830	32.34	30.698X	21.666X

Note: Reads were independently mapped from subject 412 and 416 to the assembled scaffolds from 412 or 416 to calculate sequencing depth for each subject. Only scaffolds (sequencing depth > 20 in both subjects) are listed in the table.

Table 14. Sequence statistics of metagenome vaginal samples.

Sample ID	403	405	409	412	416
Number of bases in raw metagenome	2,861,944,484	2,401,202,078	2,486,494,154	1,864,440,406	1,979,304,878
Number of bases in quality-filtered metagenome	625,577,947	466,419,347	453,049,970	198,301,156	135,032,473
Number of contigs in assembly #1	2,226	208	3,273	263	291
Total size of contigs in assembly #1 (bp)	3,238,879	121,543	6,695,385	222,561	224,266
Longest contigs length in assembly #1 (bp)	35,100	6,029	254,925	10,219	9,772
N50 contig length in assembly #1 (bp)	3,496	566	5,345	1,245	945
Number of bases aligned to assembly #1	122,155,741	19,488,753	164,582,366	27,486,163	2,326,934
Number of contigs in assembly #2	0	0	0	22	113
Total size of contigs in assembly #2 (bp)	0	0	0	122,100	133,316
Longest contigs length in assembly #2 (bp)	0	0	0	41,164	14,072
N50 contig length in assembly #2 (bp)	0	0	0	33,992	1,980
Number of bases aligned to assembly #2	0	0	0	3,369,112	2,202,156
Number of contigs in assembly #3	6,402	1,102	10,502	362	2,905
Total size of contigs in assembly #3 (bp)	7,630,506	2,843,858	9,998,666	1,530,246	2,408,033
Longest contigs length in assembly #3 (bp)	47,614	112,374	40,767	116,168	36,056
N50 contig length in assembly #3 (bp)	2,118	15,623	1,301	43,763	1,056
Number of bases aligned to assembly #3	236,831,991	298,666,171	136,881,145	76,585,271	28,288,767
Metagenome coverage	0.02	0.01	0.01	0.01	0

Note: Assemblies were composed of contigs that are >= 300 bp. Contig N50 is defined as the contig length such that using equal or longer contigs produces 50% the bases of the assembly.
Metagenome coverage is calculated as a / (b*c*d) where a = the number of bases in a quality-filtered metagenome, b = 3 Mb average bacterial genome size, c = 400 vaginal bacteria, d = 30 X sequencing depth.

Table 15. List of metabolites correlated with vaginal discharge in Yeoman *et al.* [1].

1-Acetoxy-2-Propanol
1-Hydroxy-2-Propanone
1-Methyl-4-hydroxy-1H-imidazol-2-amine
1,2-Propanediol
1,2-Propanediol,_2-acetate
1,3-Dihydro_isobenzofuran
1,4-Dimethyldioxane
2-Aminoethylphosphate
2-Ethyl-4-methyl-1,3-Dioxolane
2-Hydroxy-3-methyl-2-Cyclopenten-1-one
2-Methyl-2-hydroxybutanoic_acid
2-O-Glycerol-beta-D-galactopyranoside
2,3-Hydroxypropyl_2-aminoethylphosphate
3-Phosphoglycerate
3-Pyridinecarboxamide
Acetic_acid_(Acetate)
Alanine
Arabitol
Cyclopentanol
Ethanolamine
Gluconic_acid
Gluconic_acid-1,5-lactone
Glucose
Glutamic_acid_(Glutamate)
Glycerin
Guanine
Inositol-P
Lactic_acid_(Lactate)
N-Acetyl-serine
Ornithine
Ribose
Threonine
Valine

Table 16. Network connections among bacteria, functional genes, and metabolites.

Node 1	Node 2	Spearman correlation coefficient	FDR-adjusted p-values
2-Hydroxy-3-methyl-2-Cyclopenten-1-one	K11031 thiol-activated cytolysin	1	0
2-Ethyl-4-methyl-1,3-Dioxolane	K03320 ammonium transporter, Amt family	1	0
Gluconic_acid-1,5-lactone	K03320 ammonium transporter, Amt family	1	0
2-Hydroxy-3-methyl-2-Cyclopenten-1-one	K01000 phospho-N-acetylmuramoyl-pentapeptide-transferase	1	0
2-Hydroxy-3-methyl-2-Cyclopenten-1-one	K01736 chorismate synthase [EC:4.2.3.5]	1	0
Lactic_acid	K01262 Xaa-Pro aminopeptidase [EC:3.4.11.9]	-1	0
Lactic_acid	K01915 glutamine synthetase [EC:6.3.1.2]	-1	0
Arabitol	K03431 phosphoglucosamine mutase [EC:5.4.2.10]	-1	0
Arabitol	K00790 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	-1	0
Lactic_acid	K00133 aspartate-semialdehyde dehydrogenase [EC:1.2.1.11]	-1	0
Arabitol	K01887 arginyl-tRNA synthetase [EC:6.1.1.19]	-1	0
Arabitol	K03293 amino acid transporter, AAT family	-1	0
Lactic_acid	K00791 tRNA dimethylallyltransferase [EC:2.5.1.75]	-1	0
2-Hydroxy-3-methyl-2-Cyclopenten-1-one	K02483 two-component system, OmpR family, response regulator	1	0
Acetic_acid	K02313 chromosomal replication initiator protein	-1	0
Arabitol	K01924 UDP-N-acetylmuramate--alanine ligase [EC:6.3.2.8]	-1	0
Arabitol	K00963 UTP--glucose-1-phosphate uridylyltransferase	-1	0
2-Hydroxy-3-methyl-2-Cyclopenten-1-one	K01868 threonyl-tRNA synthetase [EC:6.1.1.3]	1	0
2-Hydroxy-3-methyl-2-Cyclopenten-1-one	K03531 cell division protein FtsZ	1	0
2-Ethyl-4-methyl-1,3-Dioxolane	K03589 cell division protein FtsQ	1	0
Ethanolamine	K03470 ribonuclease HII [EC:3.1.26.4]	1	0
Ethanolamine	K07042 probable rRNA maturation factor	1	0
Gluconic_acid-1,5-lactone	K03589 cell division protein FtsQ	1	0
Lactic_acid	K02335 DNA polymerase I [EC:2.7.7.7]	-1	0
Glutamic_acid	K02810 PTS system, sucrose-specific IIC component	1	0
Glycerin	K00599 [EC:2.1.1.-]	1	0
Arabitol	K03553 recombination protein RecA	-1	0
Arabitol	K01803 triosephosphate isomerase (TIM) [EC:5.3.1.1]	-1	0
Threonine	K00873 pyruvate kinase [EC:2.7.1.40]	-1	0
1,3-Dihydro_isobenzofuran	K08483 phosphotransferase system, enzyme I, PtsI [EC:2.7.3.9]	-1	0
2-Hydroxy-3-methyl-2-Cyclopenten-1-one	K00655 1-acyl-sn-glycerol-3-phosphate acyltransferase	1	0
Acetic_acid	K04066 primosomal protein N' (replication factor Y) (superfamily II helicase)	-1	0
Lactic_acid	K00849 galactokinase [EC:2.7.1.6]	-1	0
Ethanolamine	K03584 DNA repair protein RecO (recombination protein O)	1	0
Ethanolamine	K00806 undecaprenyl diphosphate synthase [EC:2.5.1.31]	1	0
Glutamic_acid	K02073 D-methionine transport system substrate-binding protein	1	0
Glutamic_acid	K02036 phosphate transport system ATP-binding protein	1	0
Arabitol	K09013 Fe-S cluster assembly ATP-binding protein	-1	0
Glycerin	K03647 protein involved in ribonucleotide reduction	1	0
Ethanolamine	K02503 Hit-like protein involved in cell-cycle regulation	1	0
Arabitol	K00948 ribose-phosphate pyrophosphokinase [EC:2.7.6.1]	-1	0
Threonine	K02863 large subunit ribosomal protein L1	-1	0
Glutamic_acid	K03502 DNA polymerase V	1	0
Arabitol	K06180 ribosomal large subunit pseudouridine synthase D	-1	0
Threonine	K01689 enolase [EC:4.2.1.11]	-1	0
2-Ethyl-4-methyl-1,3-Dioxolane	K03536 ribonuclease P protein component [EC:3.1.26.5]	1	0
Gluconic_acid-1,5-lactone	K03536 ribonuclease P protein component [EC:3.1.26.5]	1	0
Threonine	K01409 O-sialoglycoprotein endopeptidase [EC:3.4.24.57]	-1	0
Acetic_acid	K07738 transcriptional repressor NrdR	-1	0
Glutamic_acid	K02687 ribosomal protein L11 methyltransferase	1	0
Glycerin	K00761 uracil phosphoribosyltransferase [EC:2.4.2.9]	-1	0
1,3-Dihydro_isobenzofuran	K02933 large subunit ribosomal protein L6	-1	0
1,3-Dihydro_isobenzofuran	K02879 large subunit ribosomal protein L17	-1	0
1,3-Dihydro_isobenzofuran	K02950 small subunit ribosomal protein S12	-1	0
Glutamic_acid	K06024 segregation and condensation protein B	1	0

Table 16. (cont.)

Glutamic_acid	K05946 N-acetylglucosaminyldiphosphoundecaprenol N-acetyl-beta-D-mannosaminyltransferase	1	0
1,3-Dihydro_isobenzofuran	K02356 elongation factor P	-1	0
1,3-Dihydro_isobenzofuran	K00566 tRNA-specific 2-thiouridylase [EC:2.8.1.-]	-1	0
Glycerin	K00625 phosphate acetyltransferase [EC:2.3.1.8]	-1	0
Glycerin	K06213 magnesium transporter	-1	0
1,3-Dihydro_isobenzofuran	K11072 spermidine/putrescine transport system ATP-binding protein [EC:3.6.3.31]	-1	0
Lactic_acid	K04094 methylenetetrahydrofolate--tRNA-(uracil-5-)-methyltransferase	1	0
Lactic_acid	K03978 GTP-binding protein	1	0
Lactic_acid	K00926 carbamate kinase	1	0
2-Methyl-2-hydroxybutanoic_acid	K11752 5-amino-6-(5-phosphoribosylamino)uracil reductase	1	0
2-Methyl-2-hydroxybutanoic_acid	K00027 malate dehydrogenase (oxaloacetate-decarboxylating)	1	0
3-Pyridinecarboxamide	K00690 sucrose phosphorylase	1	0
Ornithine	K00690 sucrose phosphorylase	1	0
2-Methyl-2-hydroxybutanoic_acid	K06153 undecaprenyl-diphosphatase	1	0
K00599 [EC:2.1.1.-]	K00625 phosphate acetyltransferase [EC:2.3.1.8]	-1	0
K00761 uracil phosphoribosyltransferase	K00599 [EC:2.1.1.-]	-1	0
K00761 uracil phosphoribosyltransferase	K00625 phosphate acetyltransferase [EC:2.3.1.8]	1	0
K00791 tRNA dimethylallyltransferase	K00133 aspartate-semialdehyde dehydrogenase [EC:1.2.1.11]	1	0
K00849 galactokinase [EC:2.7.1.6]	K00133 aspartate-semialdehyde dehydrogenase [EC:1.2.1.11]	1	0
K00849 galactokinase [EC:2.7.1.6]	K00791 tRNA dimethylallyltransferase [EC:2.5.1.75]	1	0
K00926 carbamate kinase	K00133 aspartate-semialdehyde dehydrogenase [EC:1.2.1.11]	-1	0
K00926 carbamate kinase	K00791 tRNA dimethylallyltransferase [EC:2.5.1.75]	-1	0
K00926 carbamate kinase	K00849 galactokinase [EC:2.7.1.6]	-1	0
K00948 ribose-phosphate pyrophosphokinase	K00790 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	1	0
K00963 UTP--glucose-1-phosphate uridylyltransferase	K00790 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	1	0
K00963 UTP--glucose-1-phosphate uridylyltransferase	K00948 ribose-phosphate pyrophosphokinase [EC:2.7.6.1]	1	0
K01000 phospho-N-acetylmuramoyl-pentapeptide-transferase	K00655 1-acyl-sn-glycerol-3-phosphate acyltransferase	1	0
K01262 Xaa-Pro aminopeptidase	K00133 aspartate-semialdehyde dehydrogenase [EC:1.2.1.11]	1	0
K01262 Xaa-Pro aminopeptidase	K00791 tRNA dimethylallyltransferase [EC:2.5.1.75]	1	0
K01262 Xaa-Pro aminopeptidase	K00849 galactokinase [EC:2.7.1.6]	1	0
K01262 Xaa-Pro aminopeptidase	K00926 carbamate kinase	-1	0
K01409 O-sialoglycoprotein endopeptidase	K00873 pyruvate kinase [EC:2.7.1.40]	1	0
K01689 enolase [EC:4.2.1.11]	K00873 pyruvate kinase [EC:2.7.1.40]	1	0
K01689 enolase [EC:4.2.1.11]	K01409 O-sialoglycoprotein endopeptidase [EC:3.4.24.57]	1	0
K01736 chorismate synthase [EC:4.2.3.5]	K00655 1-acyl-sn-glycerol-3-phosphate acyltransferase	1	0
K01736 chorismate synthase [EC:4.2.3.5]	K01000 phospho-N-acetylmuramoyl-pentapeptide-transferase	1	0
K01803 triosephosphate isomerase (TIM)	K00790 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	1	0
K01803 triosephosphate isomerase (TIM)	K00948 ribose-phosphate pyrophosphokinase [EC:2.7.6.1]	1	0
K01803 triosephosphate isomerase (TIM)	K00963 UTP--glucose-1-phosphate uridylyltransferase	1	0
K01868 threonyl-tRNA synthetase	K00655 1-acyl-sn-glycerol-3-phosphate acyltransferase	1	0
K01868 threonyl-tRNA synthetase	K01000 phospho-N-acetylmuramoyl-pentapeptide-transferase	1	0
K01868 threonyl-tRNA synthetase	K01736 chorismate synthase [EC:4.2.3.5]	1	0
K01887 arginyl-tRNA synthetase	K00790 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	1	0
K01887 arginyl-tRNA synthetase	K00948 ribose-phosphate pyrophosphokinase [EC:2.7.6.1]	1	0
K01887 arginyl-tRNA synthetase	K00963 UTP--glucose-1-phosphate uridylyltransferase	1	0
K01887 arginyl-tRNA synthetase	K01803 triosephosphate isomerase (TIM) [EC:5.3.1.1]	1	0
K01915 glutamine synthetase [EC:6.3.1.2]	K00133 aspartate-semialdehyde dehydrogenase [EC:1.2.1.11]	1	0
K01915 glutamine synthetase [EC:6.3.1.2]	K00791 tRNA dimethylallyltransferase [EC:2.5.1.75]	1	0
K01915 glutamine synthetase [EC:6.3.1.2]	K00849 galactokinase [EC:2.7.1.6]	1	0
K01915 glutamine synthetase [EC:6.3.1.2]	K00926 carbamate kinase	-1	0
K01915 glutamine synthetase [EC:6.3.1.2]	K01262 Xaa-Pro aminopeptidase [EC:3.4.11.9]	1	0
K01924 UDP-N-acetylmuramate--alanine ligase [EC:6.3.2.8]	K00790 UDP-N-acetylglucosamine 1-carboxyvinyltransferase [EC:2.5.1.7]	1	0

Table 16. (cont.)

K01924 UDP-N-acetylmuramate--alanine ligase [EC:6.3.2.8]	K00948 ribose-phosphate pyrophosphokinase [EC:2.7.6.1]	1	0
K01924 UDP-N-acetylmuramate--alanine ligase [EC:6.3.2.8]	K00963 UTP--glucose-1-phosphate uridylyltransferase [EC:2.7.7.9]	1	0
K01924 UDP-N-acetylmuramate--alanine ligase [EC:6.3.2.8]	K01803 triosephosphate isomerase (TIM) [EC:5.3.1.1]	1	0
K01924 UDP-N-acetylmuramate--alanine ligase [EC:6.3.2.8]	K01887 arginyl-tRNA synthetase [EC:6.1.1.19]	1	0
K02073 D-methionine transport system substrate-binding protein	K02036 phosphate transport system ATP-binding protein [EC:3.6.3.27]	1	0
K00027 malate dehydrogenase (oxaloacetate-decarboxylating)	K06153 undecaprenyl-diphosphatase	1	0
K00027 malate dehydrogenase (oxaloacetate-decarboxylating)	K11752 5-amino-6-(5-phosphoribosylamino)uracil reductase	1	0
K00133 aspartate-semialdehyde dehydrogenase [EC:1.2.1.11]	K02335 DNA polymerase I [EC:2.7.7.7]	1	0
K00133 aspartate-semialdehyde dehydrogenase [EC:1.2.1.11]	K03978 GTP-binding protein	-1	0
K00133 aspartate-semialdehyde dehydrogenase [EC:1.2.1.11]	K04094 methylenetetrahydrofolate--tRNA-(uracil-5-)-methyltransferase	-1	0
K00566 tRNA-specific 2-thiouridylase	K02356 elongation factor P	1	0
K00566 tRNA-specific 2-thiouridylase	K02879 large subunit ribosomal protein L17	1	0
K00566 tRNA-specific 2-thiouridylase	K02933 large subunit ribosomal protein L6	1	0
K00566 tRNA-specific 2-thiouridylase	K02950 small subunit ribosomal protein S12	1	0
K00566 tRNA-specific 2-thiouridylase	K08483 phosphotransferase system, enzyme I, PtsI [EC:2.7.3.9]	1	0
K00566 tRNA-specific 2-thiouridylase	K11072 spermidine/putrescine transport system ATP-binding protein	1	0
K00599 [EC:2.1.1.-]	K03647 protein involved in ribonucleotide reduction	1	0
K00599 [EC:2.1.1.-]	K06213 magnesium transporter	-1	0
K00625 phosphate acetyltransferase	K03647 protein involved in ribonucleotide reduction	-1	0
K00625 phosphate acetyltransferase	K06213 magnesium transporter	1	0
K00655 1-acyl-sn-glycerol-3-phosphate acyltransferase	K02483 two-component system, OmpR family, response regulator	1	0
K00655 1-acyl-sn-glycerol-3-phosphate acyltransferase	K03531 cell division protein FtsZ	1	0
K00655 1-acyl-sn-glycerol-3-phosphate acyltransferase	K11031 thiol-activated cytolysin	1	0
K00761 uracil phosphoribosyltransferase	K03647 protein involved in ribonucleotide reduction	-1	0
K00761 uracil phosphoribosyltransferase	K06213 magnesium transporter	1	0
K00790 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	K03293 amino acid transporter, AAT family	1	0
K00790 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	K03431 phosphoglucosamine mutase [EC:5.4.2.10]	1	0
K00790 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	K03553 recombination protein RecA	1	0
K00790 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	K06180 ribosomal large subunit pseudouridine synthase D	1	0
K00790 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	K09013 Fe-S cluster assembly ATP-binding protein	1	0
K00791 tRNA dimethylallyltransferase	K02335 DNA polymerase I [EC:2.7.7.7]	1	0
K00791 tRNA dimethylallyltransferase	K03978 GTP-binding protein	-1	0
K00791 tRNA dimethylallyltransferase	K04094 methylenetetrahydrofolate--tRNA-(uracil-5-)-methyltransferase	-1	0
K00806 undecaprenyl diphosphate synthase	K02503 Hit-like protein involved in cell-cycle regulation	1	0
K00806 undecaprenyl diphosphate synthase	K03470 ribonuclease HII [EC:3.1.26.4]	1	0
K00806 undecaprenyl diphosphate synthase	K03584 DNA repair protein RecO (recombination protein O)	1	0
K00806 undecaprenyl diphosphate synthase	K07042 probable rRNA maturation factor	1	0
K00849 galactokinase [EC:2.7.1.6]	K02335 DNA polymerase I [EC:2.7.7.7]	1	0
K00849 galactokinase [EC:2.7.1.6]	K03978 GTP-binding protein	-1	0
K00849 galactokinase [EC:2.7.1.6]	K04094 methylenetetrahydrofolate--tRNA-(uracil-5-)-methyltransferase	-1	0
K00873 pyruvate kinase [EC:2.7.1.40]	K02863 large subunit ribosomal protein L1	1	0

Table 16. (cont.)

K00926 carbamate kinase	K02335 DNA polymerase I [EC:2.7.7.7]	-1	0
K00926 carbamate kinase	K03978 GTP-binding protein	1	0
K00926 carbamate kinase	K04094 methylenetetrahydrofolate--tRNA-(uracil-5-)-methyltransferase	1	0
K00948 ribose-phosphate pyrophosphokinase	K03293 amino acid transporter, AAT family	1	0
K00948 ribose-phosphate pyrophosphokinase	K03431 phosphoglucosamine mutase [EC:5.4.2.10]	1	0
K00948 ribose-phosphate pyrophosphokinase	K03553 recombination protein RecA	1	0
K00948 ribose-phosphate pyrophosphokinase	K06180 ribosomal large subunit pseudouridine synthase D	1	0
K00948 ribose-phosphate pyrophosphokinase	K09013 Fe-S cluster assembly ATP-binding protein	1	0
K00963 UTP--glucose-1-phosphate uridylyltransferase	K03293 amino acid transporter, AAT family	1	0
K00963 UTP--glucose-1-phosphate uridylyltransferase	K03431 phosphoglucosamine mutase [EC:5.4.2.10]	1	0
K00963 UTP--glucose-1-phosphate uridylyltransferase	K03553 recombination protein RecA	1	0
K00963 UTP--glucose-1-phosphate uridylyltransferase	K06180 ribosomal large subunit pseudouridine synthase D	1	0
K00963 UTP--glucose-1-phosphate uridylyltransferase	K09013 Fe-S cluster assembly ATP-binding protein	1	0
K01000 phospho-N-acetylmuramoyl-pentapeptide-transferase	K02483 two-component system, OmpR family, response regulator	1	0
K01000 phospho-N-acetylmuramoyl-pentapeptide-transferase	K03531 cell division protein FtsZ	1	0
K01000 phospho-N-acetylmuramoyl-pentapeptide-transferase	K11031 thiol-activated cytolysin	1	0
K01262 Xaa-Pro aminopeptidase	K02335 DNA polymerase I [EC:2.7.7.7]	1	0
K01262 Xaa-Pro aminopeptidase	K03978 GTP-binding protein	-1	0
K01262 Xaa-Pro aminopeptidase	K04094 methylenetetrahydrofolate--tRNA-(uracil-5-)-methyltransferase	-1	0
K01409 O-sialoglycoprotein endopeptidase	K02863 large subunit ribosomal protein L1	1	0
K01689 enolase [EC:4.2.1.11]	K02863 large subunit ribosomal protein L1	1	0
K01736 chorismate synthase [EC:4.2.3.5]	K02483 two-component system, OmpR family, response regulator	1	0
K01736 chorismate synthase [EC:4.2.3.5]	K03531 cell division protein FtsZ	1	0
K01736 chorismate synthase [EC:4.2.3.5]	K11031 thiol-activated cytolysin	1	0
K01803 triosephosphate isomerase (TIM)	K03293 amino acid transporter, AAT family	1	0
K01803 triosephosphate isomerase (TIM)	K03431 phosphoglucosamine mutase [EC:5.4.2.10]	1	0
K01803 triosephosphate isomerase (TIM)	K03553 recombination protein RecA	1	0
K01803 triosephosphate isomerase (TIM)	K06180 ribosomal large subunit pseudouridine synthase D	1	0
K01803 triosephosphate isomerase (TIM)	K09013 Fe-S cluster assembly ATP-binding protein	1	0
K01868 threonyl-tRNA synthetase	K02483 two-component system, OmpR family, response regulator	1	0
K01868 threonyl-tRNA synthetase	K03531 cell division protein FtsZ	1	0
K01868 threonyl-tRNA synthetase	K11031 thiol-activated cytolysin	1	0
K01887 arginyl-tRNA synthetase	K03293 amino acid transporter, AAT family	1	0
K01887 arginyl-tRNA synthetase	K03431 phosphoglucosamine mutase [EC:5.4.2.10]	1	0
K01887 arginyl-tRNA synthetase	K03553 recombination protein RecA	1	0
K01887 arginyl-tRNA synthetase	K06180 ribosomal large subunit pseudouridine synthase D	1	0
K01887 arginyl-tRNA synthetase	K09013 Fe-S cluster assembly ATP-binding protein	1	0
K01915 glutamine synthetase [EC:6.3.1.2]	K02335 DNA polymerase I [EC:2.7.7.7]	1	0
K01915 glutamine synthetase [EC:6.3.1.2]	K03978 GTP-binding protein	-1	0
K01915 glutamine synthetase [EC:6.3.1.2]	K04094 methylenetetrahydrofolate--tRNA-(uracil-5-)-methyltransferase	-1	0
K01924 UDP-N-acetylmuramate--alanine ligase [EC:6.3.2.8]	K03293 amino acid transporter, AAT family	1	0
K01924 UDP-N-acetylmuramate--alanine ligase [EC:6.3.2.8]	K03431 phosphoglucosamine mutase [EC:5.4.2.10]	1	0
K01924 UDP-N-acetylmuramate--alanine ligase [EC:6.3.2.8]	K03553 recombination protein RecA	1	0
K01924 UDP-N-acetylmuramate--alanine ligase [EC:6.3.2.8]	K06180 ribosomal large subunit pseudouridine synthase D	1	0
K01924 UDP-N-acetylmuramate--alanine ligase [EC:6.3.2.8]	K09013 Fe-S cluster assembly ATP-binding protein	1	0
K02036 phosphate transport system ATP-binding protein	K02687 ribosomal protein L11 methyltransferase	1	0
K02036 phosphate transport system ATP-binding protein	K02810 PTS system, sucrose-specific IIC component	1	0

Table 16. (cont.)

K02036 phosphate transport system ATP-binding protein	K03502 DNA polymerase V	1	0
K02036 phosphate transport system ATP-binding protein	K05946 N-acetylglucosaminyldiphosphoundecaprenol N-acetyl-beta-D-mannosaminyltransferase	1	0
K02036 phosphate transport system ATP-binding protein	K06024 segregation and condensation protein B	1	0
K02073 D-methionine transport system substrate-binding protein	K02687 ribosomal protein L11 methyltransferase	1	0
K02073 D-methionine transport system substrate-binding protein	K02810 PTS system, sucrose-specific IIC component	1	0
K02073 D-methionine transport system substrate-binding protein	K03502 DNA polymerase V	1	0
K02073 D-methionine transport system substrate-binding protein	K05946 N-acetylglucosaminyldiphosphoundecaprenol N-acetyl-beta-D-mannosaminyltransferase	1	0
K02073 D-methionine transport system substrate-binding protein	K06024 segregation and condensation protein B	1	0
K02313 chromosomal replication initiator protein	K04066 primosomal protein N' (replication factor Y) (superfamily II helicase) [EC:3.6.4.-]	1	0
K02313 chromosomal replication initiator protein	K07738 transcriptional repressor NrdR	1	0
K02335 DNA polymerase I [EC:2.7.7.7]	K03978 GTP-binding protein	-1	0
K02335 DNA polymerase I [EC:2.7.7.7]	K04094 methylenetetrahydrofolate--tRNA-(uracil-5-)-methyltransferase	-1	0
K02356 elongation factor P	K02879 large subunit ribosomal protein L17	1	0
K02356 elongation factor P	K02933 large subunit ribosomal protein L6	1	0
K02356 elongation factor P	K02950 small subunit ribosomal protein S12	1	0
K02356 elongation factor P	K08483 phosphotransferase system, enzyme I, PtsI [EC:2.7.3.9]	1	0
K02356 elongation factor P	K11072 spermidine/putrescine transport system ATP-binding protein	1	0
K02483 two-component system, OmpR family, response regulator	K03531 cell division protein FtsZ	1	0
K02483 two-component system, OmpR family, response regulator	K11031 thiol-activated cytolysin	1	0
K02503 Hit-like protein involved in cell-cycle regulation	K03470 ribonuclease HII [EC:3.1.26.4]	1	0
K02503 Hit-like protein involved in cell-cycle regulation	K03584 DNA repair protein RecO (recombination protein O)	1	0
K02503 Hit-like protein involved in cell-cycle regulation	K07042 probable rRNA maturation factor	1	0
K02687 ribosomal protein L11 methyltransferase	K02810 PTS system, sucrose-specific IIC component	1	0
K02687 ribosomal protein L11 methyltransferase	K03502 DNA polymerase V	1	0
K02687 ribosomal protein L11 methyltransferase	K05946 N-acetylglucosaminyldiphosphoundecaprenol N-acetyl-beta-D-mannosaminyltransferase	1	0
K02687 ribosomal protein L11 methyltransferase	K06024 segregation and condensation protein B	1	0
K02810 PTS system, sucrose-specific IIC component	K03502 DNA polymerase V	1	0
K02810 PTS system, sucrose-specific IIC component	K05946 N-acetylglucosaminyldiphosphoundecaprenol N-acetyl-beta-D-mannosaminyltransferase	1	0
K02810 PTS system, sucrose-specific IIC component	K06024 segregation and condensation protein B	1	0
K02879 large subunit ribosomal protein L17	K02933 large subunit ribosomal protein L6	1	0
K02879 large subunit ribosomal protein L17	K02950 small subunit ribosomal protein S12	1	0
K02879 large subunit ribosomal protein L17	K08483 phosphotransferase system, enzyme I, PtsI [EC:2.7.3.9]	1	0
K02879 large subunit ribosomal protein L17	K11072 spermidine/putrescine transport system ATP-binding protein	1	0
K02933 large subunit ribosomal protein L6	K02950 small subunit ribosomal protein S12	1	0
K02933 large subunit ribosomal protein L6	K08483 phosphotransferase system, enzyme I, PtsI [EC:2.7.3.9]	1	0
K02933 large subunit ribosomal protein L6	K11072 spermidine/putrescine transport system ATP-binding protein	1	0
K02950 small subunit ribosomal protein S12	K08483 phosphotransferase system, enzyme I, PtsI [EC:2.7.3.9]	1	0
K02950 small subunit ribosomal protein S12	K11072 spermidine/putrescine transport system ATP-binding protein	1	0
K03293 amino acid transporter, AAT family	K03431 phosphoglucosamine mutase [EC:5.4.2.10]	1	0
K03293 amino acid transporter, AAT family	K03553 recombination protein RecA	1	0
K03293 amino acid transporter, AAT family	K06180 ribosomal large subunit pseudouridine synthase D	1	0

Table 16. (cont.)

K03293 amino acid transporter, AAT family	K09013 Fe-S cluster assembly ATP-binding protein	1	0
K03320 ammonium transporter, Amt family	K03536 ribonuclease P protein component [EC:3.1.26.5]	1	0
K03320 ammonium transporter, Amt family	K03589 cell division protein FtsQ	1	0
K03431 phosphoglucosamine mutase	K03553 recombination protein RecA	1	0
K03431 phosphoglucosamine mutase	K06180 ribosomal large subunit pseudouridine synthase D	1	0
K03431 phosphoglucosamine mutase	K09013 Fe-S cluster assembly ATP-binding protein	1	0
K03470 ribonuclease HII [EC:3.1.26.4]	K03584 DNA repair protein RecO (recombination protein O)	1	0
K03470 ribonuclease HII [EC:3.1.26.4]	K07042 probable rRNA maturation factor	1	0
K03502 DNA polymerase V	K05946 N-acetylglucosaminyldiphosphoundecaprenol N-acetyl-beta-D-mannosaminyltransferase	1	0
K03502 DNA polymerase V	K06024 segregation and condensation protein B	1	0
K11031 thiol-activated cytolysin	K03531 cell division protein FtsZ	1	0
K03589 cell division protein FtsQ	K03536 ribonuclease P protein component [EC:3.1.26.5]	1	0
K03553 recombination protein RecA	K06180 ribosomal large subunit pseudouridine synthase D	1	0
K03553 recombination protein RecA	K09013 Fe-S cluster assembly ATP-binding protein	1	0
K07042 probable rRNA maturation factor	K03584 DNA repair protein RecO (recombination protein O)	1	0
K06213 magnesium transporter	K03647 protein involved in ribonucleotide reduction	-1	0
K04094 methylenetetrahydrofolate--tRNA-(uracil-5-)-methyltransferase	K03978 GTP-binding protein	1	0
K07738 transcriptional repressor NrdR	K04066 primosomal protein N' (replication factor Y) (superfamily II helicase) [EC:3.6.4.-]	1	0
K06024 segregation and condensation protein B	K05946 N-acetylglucosaminyldiphosphoundecaprenol N-acetyl-beta-D-mannosaminyltransferase	1	0
K11752 5-amino-6-(5-phosphoribosylamino)uracil reductase	K06153 undecaprenyl-diphosphatase	1	0
K09013 Fe-S cluster assembly ATP-binding protein	K06180 ribosomal large subunit pseudouridine synthase D	1	0
K11072 spermidine/putrescine transport system ATP-binding protein	K08483 phosphotransferase system, enzyme I, PtsI [EC:2.7.3.9]	1	0
2-Methyl-2-hydroxybutanoic_acid	Aquabacterium	1	0
2-Methyl-2-hydroxybutanoic_acid	Chryseobacterium	1	0
2-Methyl-2-hydroxybutanoic_acid	Elizabethkingia	1	0
2-Methyl-2-hydroxybutanoic_acid	Finegoldia	1	0
2-Methyl-2-hydroxybutanoic_acid	Morganella	1	0
2-Methyl-2-hydroxybutanoic_acid	Negativicoccus	1	0
2-Methyl-2-hydroxybutanoic_acid	Phenylobacterium	1	0
2-Methyl-2-hydroxybutanoic_acid	Propionibacterium	1	0
2-Methyl-2-hydroxybutanoic_acid	Serratia	1	0
2-Methyl-2-hydroxybutanoic_acid	Siphonobacter	1	0
2-Hydroxy-3-methyl-2-Cyclopenten-1-one	Bulleidia	1	0
1,2-Propanediol	Peptoniphilus	1	0
3-Pyridinecarboxamide	Chlorophyta	1	0
Glucose	Janthinobacterium	-1	0
Ornithine	Chlorophyta	1	0
Ribose	Acinetobacter	1	0
Valine	Ureaplasma	-1	0
K00027 malate dehydrogenase (oxaloacetate-decarboxylating)	Aquabacterium	1	0
K00027 malate dehydrogenase (oxaloacetate-decarboxylating)	Chryseobacterium	1	0
K00027 malate dehydrogenase (oxaloacetate-decarboxylating)	Elizabethkingia	1	0
K00027 malate dehydrogenase (oxaloacetate-decarboxylating)	Finegoldia	1	0
K00027 malate dehydrogenase (oxaloacetate-decarboxylating)	Morganella	1	0
K00027 malate dehydrogenase (oxaloacetate-decarboxylating)	Negativicoccus	1	0
K00027 malate dehydrogenase (oxaloacetate-decarboxylating)	Phenylobacterium	1	0
K00027 malate dehydrogenase (oxaloacetate-decarboxylating)	Propionibacterium	1	0

Table 16. (cont.)

K00027 malate dehydrogenase (oxaloacetate-decarboxylating)	Serratia	1	0
K00027 malate dehydrogenase (oxaloacetate-decarboxylating)	Siphonobacter	1	0
K00655 1-acyl-sn-glycerol-3-phosphate acyltransferase [EC:2.3.1.51]	Bulleidia	1	0
K00690 sucrose phosphorylase	Chlorophyta	1	0
K01000 phospho-N-acetylmuramoyl-pentapeptide-transferase	Bulleidia	1	0
K01736 chorismate synthase [EC:4.2.3.5]	Bulleidia	1	0
K01868 threonyl-tRNA synthetase	Bulleidia	1	0
K02483 two-component system, OmpR family, response regulator	Bulleidia	1	0
K03531 cell division protein FtsZ	Bulleidia	1	0
K06153 undecaprenyl-diphosphatase	Aquabacterium	1	0
K06153 undecaprenyl-diphosphatase	Chryseobacterium	1	0
K06153 undecaprenyl-diphosphatase	Elizabethkingia	1	0
K06153 undecaprenyl-diphosphatase	Finegoldia	1	0
K06153 undecaprenyl-diphosphatase	Morganella	1	0
K06153 undecaprenyl-diphosphatase	Negativicoccus	1	0
K06153 undecaprenyl-diphosphatase	Phenylobacterium	1	0
K06153 undecaprenyl-diphosphatase	Propionibacterium	1	0
K06153 undecaprenyl-diphosphatase	Serratia	1	0
K06153 undecaprenyl-diphosphatase	Siphonobacter	1	0
K11752 5-amino-6-(5-phosphoribosylamino)uracil reductase	Aquabacterium	1	0
K11752 5-amino-6-(5-phosphoribosylamino)uracil reductase	Chryseobacterium	1	0
K11752 5-amino-6-(5-phosphoribosylamino)uracil reductase	Elizabethkingia	1	0
K11752 5-amino-6-(5-phosphoribosylamino)uracil reductase	Finegoldia	1	0
K11752 5-amino-6-(5-phosphoribosylamino)uracil reductase	Morganella	1	0
K11752 5-amino-6-(5-phosphoribosylamino)uracil reductase	Negativicoccus	1	0
K11752 5-amino-6-(5-phosphoribosylamino)uracil reductase	Phenylobacterium	1	0
K11752 5-amino-6-(5-phosphoribosylamino)uracil reductase	Propionibacterium	1	0
K11752 5-amino-6-(5-phosphoribosylamino)uracil reductase	Serratia	1	0
K11752 5-amino-6-(5-phosphoribosylamino)uracil reductase	Siphonobacter	1	0
K11031 thiol-activated cytolysin	Bulleidia	1	0
Chryseobacterium	Aquabacterium	1	0
Elizabethkingia	Aquabacterium	1	0
Elizabethkingia	Chryseobacterium	1	0
Finegoldia	Aquabacterium	1	0
Finegoldia	Chryseobacterium	1	0
Finegoldia	Elizabethkingia	1	0
Morganella	Aquabacterium	1	0
Negativicoccus	Aquabacterium	1	0
Phenylobacterium	Aquabacterium	1	0
Propionibacterium	Aquabacterium	1	0
Serratia	Aquabacterium	1	0
Siphonobacter	Aquabacterium	1	0
Morganella	Chryseobacterium	1	0
Negativicoccus	Chryseobacterium	1	0
Phenylobacterium	Chryseobacterium	1	0
Propionibacterium	Chryseobacterium	1	0
Serratia	Chryseobacterium	1	0

Table 16. (cont.)

Siphonobacter	Chryseobacterium	1	0
Morganella	Elizabethkingia	1	0
Negativicoccus	Elizabethkingia	1	0
Phenylobacterium	Elizabethkingia	1	0
Propionibacterium	Elizabethkingia	1	0
Serratia	Elizabethkingia	1	0
Siphonobacter	Elizabethkingia	1	0
Morganella	Finegoldia	1	0
Negativicoccus	Finegoldia	1	0
Phenylobacterium	Finegoldia	1	0
Propionibacterium	Finegoldia	1	0
Serratia	Finegoldia	1	0
Siphonobacter	Finegoldia	1	0
Negativicoccus	Morganella	1	0
Phenylobacterium	Morganella	1	0
Propionibacterium	Morganella	1	0
Serratia	Morganella	1	0
Siphonobacter	Morganella	1	0
Phenylobacterium	Negativicoccus	1	0
Propionibacterium	Negativicoccus	1	0
Serratia	Negativicoccus	1	0
Siphonobacter	Negativicoccus	1	0
Propionibacterium	Phenylobacterium	1	0
Serratia	Phenylobacterium	1	0
Siphonobacter	Phenylobacterium	1	0
Serratia	Propionibacterium	1	0
Siphonobacter	Propionibacterium	1	0
Siphonobacter	Serratia	1	0

Table 17. Taxonomic annotations of functional genes correlated with discharge-associated metabolites.

KO ID	KO Description	Open reading frames ID	Name of bacterial strain
K00027	malate dehydrogenase (oxaloacetate-decarboxylating)	416_2237_2237_1	Gardnerella vaginalis 409-05
K00133	aspartate-semialdehyde dehydrogenase	403_976_976_21	Gardnerella vaginalis 409-05
K00133	aspartate-semialdehyde dehydrogenase	403_2578_2578_3	Gardnerella vaginalis 409-05
K00133	aspartate-semialdehyde dehydrogenase	403_7327_7327_11	Gardnerella vaginalis ATCC 14018 = JCM 11026
K00133	aspartate-semialdehyde dehydrogenase	403_2588_2588_4	Gardnerella vaginalis ATCC 14018 = JCM 11026
K00133	aspartate-semialdehyde dehydrogenase	405_252_252_33	Gardnerella vaginalis ATCC 14018 = JCM 11026
K00133	aspartate-semialdehyde dehydrogenase	409_1955_1955_1	Gardnerella vaginalis ATCC 14018 = JCM 11026
K00133	aspartate-semialdehyde dehydrogenase	409_2728_2728_1	Gardnerella vaginalis ATCC 14018 = JCM 11026
K00566	tRNA-specific 2-thiouridylase	409_12116_12116_1	Gardnerella vaginalis ATCC 14018 = JCM 11026
K00566	tRNA-specific 2-thiouridylase	416_2369_2369_1	Gardnerella vaginalis HMP9231
K00566	tRNA-specific 2-thiouridylase	403_2996_2996_1	Lactobacillus acetotolerans
K00566	tRNA-specific 2-thiouridylase	405_375_375_24	Lactobacillus acidophilus strain FSI4
K00566	tRNA-specific 2-thiouridylase	409_8276_8276_1	Lactobacillus acidophilus strain FSI4
K00566	tRNA-specific 2-thiouridylase	409_8138_8138_1	Lactobacillus acidophilus strain FSI4
K00599	[EC:2.1.1.-]	403_797_797_1	Lactobacillus acidophilus strain FSI4
K00599	[EC:2.1.1.-]	403_1632_1632_2	Lactobacillus acidophilus strain FSI4
K00599	[EC:2.1.1.-]	409_581_581_199	Lactobacillus delbrueckii subsp. bulgaricus ND02
K00599	[EC:2.1.1.-]	405_386_386_8	Lactobacillus delbrueckii subsp. bulgaricus ND02
K00599	[EC:2.1.1.-]	403_2626_2626_4	Lactobacillus gasseri 130918
K00599	[EC:2.1.1.-]	403_2833_2833_3	Lactobacillus gasseri 130918
K00599	[EC:2.1.1.-]	405_247_247_29	Lactobacillus gasseri 130918
K00599	[EC:2.1.1.-]	409_3006_3006_1	Lactobacillus gasseri 130918
K00625	phosphate acetyltransferase	405_271_271_5	Lactobacillus gasseri ATCC 33323
K00625	phosphate acetyltransferase	403_3042_3042_2	Lactobacillus gasseri ATCC 33323
K00625	phosphate acetyltransferase	409_1889_1889_1	Lactobacillus gasseri ATCC 33323
K00625	phosphate acetyltransferase	409_2609_2609_2	Lactobacillus gasseri ATCC 33323
K00625	phosphate acetyltransferase	403_5931_5931_1	Lactobacillus gasseri ATCC 33323
K00625	phosphate acetyltransferase	403_6927_6927_1	Lactobacillus helveticus H9
K00625	phosphate acetyltransferase	409_6597_6597_1	Lactobacillus helveticus strain MB2-1
K00655	1-acyl-sn-glycerol-3-phosphate acyltransferase	403_1130_1130_11	Lactobacillus helveticus strain MB2-1
K00655	1-acyl-sn-glycerol-3-phosphate acyltransferase	403_2285_2285_7	Lactobacillus helveticus strain MB2-1
K00655	1-acyl-sn-glycerol-3-phosphate acyltransferase	403_3322_3322_3	Lactobacillus johnsonii FI9785
K00655	1-acyl-sn-glycerol-3-phosphate acyltransferase	403_2640_2640_3	Lactobacillus johnsonii FI9785
K00655	1-acyl-sn-glycerol-3-phosphate acyltransferase	405_1310_1310_105	Lactobacillus johnsonii N6.2
K00655	1-acyl-sn-glycerol-3-phosphate acyltransferase	409_2324_2324_7	Lactobacillus johnsonii N6.2
K00690	sucrose phosphorylase	405_430_430_1	Lactobacillus johnsonii N6.2
K00761	uracil phosphoribosyltransferase	403_957_957_14	Lactobacillus johnsonii N6.2
K00761	uracil phosphoribosyltransferase	403_3712_3712_1	Lactobacillus johnsonii N6.2
K00761	uracil phosphoribosyltransferase	412_379_379_20	Lactobacillus johnsonii N6.2
K00761	uracil phosphoribosyltransferase	405_1299_1299_27	Lactobacillus kefiranofaciens ZW3
K00761	uracil phosphoribosyltransferase	403_126_126_1	Lactobacillus kefiranofaciens ZW3
K00761	uracil phosphoribosyltransferase	409_2503_2503_2	Lactobacillus kimbladii strain Hma2N
K00761	uracil phosphoribosyltransferase	409_4777_4777_1	Lactobacillus sp. wkB8
K00761	uracil phosphoribosyltransferase	409_3706_3706_2	Mycoplasma hominis strain Sprott
K11031	thiol-activated cytolysin	403_8310_8310_4	Prevotella intermedia 17
K11031	thiol-activated cytolysin	409_574_574_8	Prevotella melaninogenica ATCC 25845
K11072	spermidine/putrescine transport system ATP-binding protein	405_250_250_17	Pseudomonas chlororaphis strain PA23
K11072	spermidine/putrescine transport system ATP-binding protein	416_1729_1729_1	Pseudomonas chlororaphis strain PA23
K11072	spermidine/putrescine transport system ATP-binding protein	403_1946_1946_1	Pseudomonas fluorescens SBW25
K11072	spermidine/putrescine transport system ATP-binding protein	403_6499_6499_1	Pseudomonas rhizosphaerae strain DSM 16299
K11072	spermidine/putrescine transport system ATP-binding protein	409_3876_3876_2	Streptobacillus moniliformis DSM 12112
K11072	spermidine/putrescine transport system ATP-binding protein	416_2437_2437_1	Ureaplasma parvum serovar 3
K11072	spermidine/putrescine transport system ATP-binding protein	409_10674_10674_1	Ureaplasma urealyticum serovar 10
K11072	spermidine/putrescine transport system ATP-binding protein	409_11047_11047_1	Ureaplasma urealyticum serovar 10
K11072	spermidine/putrescine transport system ATP-binding protein	409_8054_8054_1	Ureaplasma urealyticum serovar 10 str. ATCC 33699
K11752	5-amino-6-(5-phosphoribosylamino)uracil reductase	416_2289_2289_1	Veillonella parvula DSM 2008
K00790	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	403_995_995_3	NA
K00790	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	403_2949_2949_4	NA
K00790	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	403_3001_3001_2	NA
K00790	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	403_2774_2774_2	NA
K00790	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	405_862_862_11	NA
K00790	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	416_917_917_1	NA
K00790	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	409_824_824_2	NA

Table 17. (cont.)

K00790	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	409_2088_2088_1	NA
K00790	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	409_10581_10581_1	NA
K00791	tRNA dimethylallyltransferase	403_1238_1238_1	NA
K00791	tRNA dimethylallyltransferase	403_2374_2374_5	NA
K00791	tRNA dimethylallyltransferase	403_2414_2414_2	NA
K00791	tRNA dimethylallyltransferase	405_267_267_24	NA
K00791	tRNA dimethylallyltransferase	403_5148_5148_1	NA
K00806	undecaprenyl diphosphate synthase	403_1063_1063_8	NA
K00806	undecaprenyl diphosphate synthase	405_1310_1310_99	NA
K00806	undecaprenyl diphosphate synthase	403_2604_2604_5	NA
K00806	undecaprenyl diphosphate synthase	403_3010_3010_1	NA
K00849	galactokinase	403_2598_2598_8	NA
K00849	galactokinase	403_2726_2726_3	NA
K00849	galactokinase	409_1442_1442_1	NA
K00849	galactokinase	409_2898_2898_2	NA
K00873	pyruvate kinase	403_1110_1110_2	NA
K00873	pyruvate kinase	403_8044_8044_2	NA
K00873	pyruvate kinase	405_242_242_12	NA
K00873	pyruvate kinase	403_2952_2952_2	NA
K00873	pyruvate kinase	403_3177_3177_2	NA
K00873	pyruvate kinase	412_331_331_56	NA
K00873	pyruvate kinase	403_38_38_1	NA
K00873	pyruvate kinase	409_2359_2359_1	NA
K00926	carbamate kinase	405_258_258_6	NA
K00926	carbamate kinase	409_8028_8028_1	NA
K00948	ribose-phosphate pyrophosphokinase	403_2270_2270_1	NA
K00948	ribose-phosphate pyrophosphokinase	403_4073_4073_3	NA
K00948	ribose-phosphate pyrophosphokinase	409_609_609_6	NA
K00948	ribose-phosphate pyrophosphokinase	403_1260_1260_1	NA
K00948	ribose-phosphate pyrophosphokinase	403_2575_2575_5	NA
K00948	ribose-phosphate pyrophosphokinase	403_8210_8210_1	NA
K00948	ribose-phosphate pyrophosphokinase	405_282_282_2	NA
K00948	ribose-phosphate pyrophosphokinase	416_744_744_1	NA
K00948	ribose-phosphate pyrophosphokinase	405_627_627_1	NA
K00948	ribose-phosphate pyrophosphokinase	405_1240_1240_1	NA
K00948	ribose-phosphate pyrophosphokinase	416_1595_1595_1	NA
K00948	ribose-phosphate pyrophosphokinase	403_8623_8623_13	NA
K00948	ribose-phosphate pyrophosphokinase	409_3696_3696_3	NA
K00948	ribose-phosphate pyrophosphokinase	403_2254_2254_11	NA
K00948	ribose-phosphate pyrophosphokinase	405_1242_1242_1	NA
K00948	ribose-phosphate pyrophosphokinase	412_341_341_36	NA
K00948	ribose-phosphate pyrophosphokinase	416_3135_3135_1	NA
K00948	ribose-phosphate pyrophosphokinase	403_367_367_4	NA
K00948	ribose-phosphate pyrophosphokinase	409_1604_1604_1	NA
K00948	ribose-phosphate pyrophosphokinase	409_2874_2874_3	NA
K00948	ribose-phosphate pyrophosphokinase	409_11574_11574_1	NA
K00948	ribose-phosphate pyrophosphokinase	409_11799_11799_1	NA
K00963	UTP--glucose-1-phosphate uridylyltransferase	403_1150_1150_10	NA
K00963	UTP--glucose-1-phosphate uridylyltransferase	403_2291_2291_3	NA
K00963	UTP--glucose-1-phosphate uridylyltransferase	403_8618_8618_6	NA
K00963	UTP--glucose-1-phosphate uridylyltransferase	416_545_545_1	NA
K00963	UTP--glucose-1-phosphate uridylyltransferase	403_3482_3482_1	NA
K00963	UTP--glucose-1-phosphate uridylyltransferase	403_4795_4795_1	NA
K00963	UTP--glucose-1-phosphate uridylyltransferase	405_318_318_5	NA
K00963	UTP--glucose-1-phosphate uridylyltransferase	412_303_303_36	NA
K01000	phospho-N-acetylmuramoyl-pentapeptide-transferase	403_1239_1239_10	NA
K01000	phospho-N-acetylmuramoyl-pentapeptide-transferase	403_2363_2363_6	NA
K01000	phospho-N-acetylmuramoyl-pentapeptide-transferase	403_2908_2908_21	NA
K01000	phospho-N-acetylmuramoyl-pentapeptide-transferase	405_375_375_7	NA
K01000	phospho-N-acetylmuramoyl-pentapeptide-transferase	403_2615_2615_5	NA
K01262	Xaa-Pro aminopeptidase	403_1064_1064_10	NA
K01262	Xaa-Pro aminopeptidase	403_2368_2368_4	NA
K01262	Xaa-Pro aminopeptidase	403_8346_8346_2	NA
K01262	Xaa-Pro aminopeptidase	405_264_264_33	NA
K01262	Xaa-Pro aminopeptidase	403_3998_3998_1	NA
K01262	Xaa-Pro aminopeptidase	409_1819_1819_2	NA
K01262	Xaa-Pro aminopeptidase	409_2801_2801_1	NA

Table 17. (cont.)

K01409	O-sialoglycoprotein endopeptidase	403_1189_1189_1	NA
K01409	O-sialoglycoprotein endopeptidase	403_8369_8369_3	NA
K01409	O-sialoglycoprotein endopeptidase	403_2606_2606_7	NA
K01409	O-sialoglycoprotein endopeptidase	405_263_263_23	NA
K01409	O-sialoglycoprotein endopeptidase	416_722_722_5	NA
K01409	O-sialoglycoprotein endopeptidase	416_2554_2554_1	NA
K01409	O-sialoglycoprotein endopeptidase	409_835_835_2	NA
K01409	O-sialoglycoprotein endopeptidase	409_2470_2470_2	NA
K01409	O-sialoglycoprotein endopeptidase	409_2471_2471_1	NA
K01689	enolase	403_2984_2984_1	NA
K01689	enolase	403_3327_3327_1	NA
K01689	enolase	409_586_586_108	NA
K01689	enolase	403_3442_3442_2	NA
K01689	enolase	403_4869_4869_1	NA
K01689	enolase	412_287_287_8	NA
K01689	enolase	416_1285_1285_1	NA
K01689	enolase	403_2586_2586_14	NA
K01689	enolase	403_2586_2586_15	NA
K01689	enolase	403_2584_2584_8	NA
K01689	enolase	405_290_290_3	NA
K01689	enolase	412_430_430_27	NA
K01689	enolase	409_54_54_2	NA
K01689	enolase	409_1226_1226_6	NA
K01689	enolase	409_2124_2124_1	NA
K01689	enolase	409_2125_2125_1	NA
K01689	enolase	409_3565_3565_1	NA
K01736	chorismate synthase	403_1025_1025_5	NA
K01736	chorismate synthase	403_7606_7606_1	NA
K01736	chorismate synthase	403_8205_8205_1	NA
K01803	triosephosphate isomerase (TIM)	403_1177_1177_2	NA
K01803	triosephosphate isomerase (TIM)	403_2344_2344_8	NA
K01803	triosephosphate isomerase (TIM)	403_7853_7853_1	NA
K01803	triosephosphate isomerase (TIM)	403_3442_3442_1	NA
K01803	triosephosphate isomerase (TIM)	405_331_331_2	NA
K01803	triosephosphate isomerase (TIM)	412_287_287_9	NA
K01803	triosephosphate isomerase (TIM)	409_2232_2232_7	NA
K01868	threonyl-tRNA synthetase	403_1161_1161_10	NA
K01868	threonyl-tRNA synthetase	403_8268_8268_1	NA
K01868	threonyl-tRNA synthetase	403_1429_1429_1	NA
K01868	threonyl-tRNA synthetase	403_8075_8075_1	NA
K01868	threonyl-tRNA synthetase	403_2636_2636_4	NA
K01868	threonyl-tRNA synthetase	405_278_278_34	NA
K01868	threonyl-tRNA synthetase	409_1389_1389_7	NA
K01868	threonyl-tRNA synthetase	409_1390_1390_1	NA
K01868	threonyl-tRNA synthetase	409_1391_1391_1	NA
K01868	threonyl-tRNA synthetase	409_2180_2180_3	NA
K01868	threonyl-tRNA synthetase	409_13227_13227_1	NA
K01868	threonyl-tRNA synthetase	409_4102_4102_1	NA
K01887	arginyl-tRNA synthetase	403_2572_2572_2	NA
K01887	arginyl-tRNA synthetase	403_8409_8409_2	NA
K01887	arginyl-tRNA synthetase	405_337_337_7	NA
K01887	arginyl-tRNA synthetase	403_2939_2939_1	NA
K01887	arginyl-tRNA synthetase	412_299_299_7	NA
K01887	arginyl-tRNA synthetase	409_1247_1247_2	NA
K01887	arginyl-tRNA synthetase	409_2609_2609_7	NA
K01915	glutamine synthetase	403_1241_1241_1	NA
K01915	glutamine synthetase	403_2320_2320_1	NA
K01915	glutamine synthetase	403_8225_8225_10	NA
K01915	glutamine synthetase	409_4805_4805_2	NA
K01915	glutamine synthetase	409_11332_11332_1	NA
K01915	glutamine synthetase	403_1248_1248_3	NA
K01915	glutamine synthetase	403_2421_2421_1	NA
K01915	glutamine synthetase	403_8200_8200_3	NA
K01915	glutamine synthetase	405_267_267_22	NA
K01915	glutamine synthetase	403_3164_3164_2	NA
K01915	glutamine synthetase	405_1259_1259_1	NA

Table 17. (cont.)

K01915	glutamine synthetase	412_430_430_20	NA
K01915	glutamine synthetase	416_1346_1346_1	NA
K01915	glutamine synthetase	409_6530_6530_2	NA
K01915	glutamine synthetase	403_5453_5453_1	NA
K01915	glutamine synthetase	403_5091_5091_1	NA
K01915	glutamine synthetase	409_8789_8789_1	NA
K01924	UDP-N-acetylmuramate--alanine ligase	403_1239_1239_6	NA
K01924	UDP-N-acetylmuramate--alanine ligase	403_2363_2363_2	NA
K01924	UDP-N-acetylmuramate--alanine ligase	403_3421_3421_1	NA
K01924	UDP-N-acetylmuramate--alanine ligase	405_278_278_20	NA
K01924	UDP-N-acetylmuramate--alanine ligase	416_1590_1590_1	NA
K02036	phosphate transport system ATP-binding protein	403_999_999_1	NA
K02036	phosphate transport system ATP-binding protein	403_2472_2472_2	NA
K02036	phosphate transport system ATP-binding protein	405_262_262_2	NA
K02036	phosphate transport system ATP-binding protein	405_215_215_11	NA
K02036	phosphate transport system ATP-binding protein	405_262_262_3	NA
K02036	phosphate transport system ATP-binding protein	403_2597_2597_4	NA
K02036	phosphate transport system ATP-binding protein	403_2597_2597_3	NA
K02036	phosphate transport system ATP-binding protein	403_4830_4830_3	NA
K02073	D-methionine transport system substrate-binding protein	403_1168_1168_4	NA
K02073	D-methionine transport system substrate-binding protein	403_4044_4044_1	NA
K02073	D-methionine transport system substrate-binding protein	405_673_673_7	NA
K02073	D-methionine transport system substrate-binding protein	405_236_236_10	NA
K02073	D-methionine transport system substrate-binding protein	403_2584_2584_2	NA
K02073	D-methionine transport system substrate-binding protein	403_5136_5136_1	NA
K02313	chromosomal replication initiator protein	403_954_954_20	NA
K02313	chromosomal replication initiator protein	403_8591_8591_3	NA
K02313	chromosomal replication initiator protein	405_211_211_9	NA
K02313	chromosomal replication initiator protein	403_8598_8598_9	NA
K02313	chromosomal replication initiator protein	412_344_344_23	NA
K02313	chromosomal replication initiator protein	409_2128_2128_1	NA
K02335	DNA polymerase I	403_1115_1115_3	NA
K02335	DNA polymerase I	403_2620_2620_3	NA
K02335	DNA polymerase I	405_278_278_28	NA
K02335	DNA polymerase I	409_3313_3313_1	NA
K02335	DNA polymerase I	403_4490_4490_1	NA
K02356	elongation factor P	403_2947_2947_5	NA
K02356	elongation factor P	405_427_427_10	NA
K02356	elongation factor P	405_264_264_32	NA
K02356	elongation factor P	403_3998_3998_2	NA
K02356	elongation factor P	403_1852_1852_1	NA
K02356	elongation factor P	403_2966_2966_1	NA
K02356	elongation factor P	409_3661_3661_2	NA
K02356	elongation factor P	403_449_449_6	NA
K02356	elongation factor P	409_1435_1435_2	NA
K02483	two-component system, OmpR family, response regulator	403_1134_1134_24	NA
K02483	two-component system, OmpR family, response regulator	403_2923_2923_5	NA
K02483	two-component system, OmpR family, response regulator	403_3017_3017_1	NA
K02503	Hit-like protein involved in cell-cycle regulation	403_1165_1165_14	NA
K02503	Hit-like protein involved in cell-cycle regulation	403_2540_2540_7	NA
K02503	Hit-like protein involved in cell-cycle regulation	405_278_278_13	NA
K02503	Hit-like protein involved in cell-cycle regulation	409_2564_2564_3	NA
K02687	ribosomal protein L11 methyltransferase	403_2659_2659_3	NA
K02687	ribosomal protein L11 methyltransferase	405_238_238_19	NA
K02810	PTS system, sucrose-specific IIC component	405_738_738_7	NA
K02810	PTS system, sucrose-specific IIC component	403_2651_2651_6	NA
K02810	PTS system, sucrose-specific IIC component	403_1878_1878_1	NA
K02863	large subunit ribosomal protein L1	403_1286_1286_2	NA
K02863	large subunit ribosomal protein L1	403_2249_2249_3	NA
K02863	large subunit ribosomal protein L1	405_262_262_7	NA
K02863	large subunit ribosomal protein L1	403_3250_3250_1	NA
K02863	large subunit ribosomal protein L1	403_337_337_4	NA
K02879	large subunit ribosomal protein L17	403_1001_1001_27	NA
K02879	large subunit ribosomal protein L17	409_3783_3783_2	NA
K02879	large subunit ribosomal protein L17	403_2600_2600_7	NA
K02879	large subunit ribosomal protein L17	405_235_235_8	NA
K02879	large subunit ribosomal protein L17	412_341_341_8	NA

Table 17. (cont.)

K02879	large subunit ribosomal protein L17	403_359_359_6	NA
K02933	large subunit ribosomal protein L6	403_1001_1001_16	NA
K02933	large subunit ribosomal protein L6	403_2406_2406_6	NA
K02933	large subunit ribosomal protein L6	403_5114_5114_2	NA
K02933	large subunit ribosomal protein L6	412_341_341_19	NA
K02933	large subunit ribosomal protein L6	405_235_235_20	NA
K02933	large subunit ribosomal protein L6	403_356_356_3	NA
K02933	large subunit ribosomal protein L6	409_2049_2049_1	NA
K02950	small subunit ribosomal protein S12	403_342_342_1	NA
K02950	small subunit ribosomal protein S12	403_1239_1239_4	NA
K02950	small subunit ribosomal protein S12	403_2431_2431_3	NA
K02950	small subunit ribosomal protein S12	409_4214_4214_2	NA
K02950	small subunit ribosomal protein S12	403_5044_5044_1	NA
K02950	small subunit ribosomal protein S12	405_369_369_6	NA
K02950	small subunit ribosomal protein S12	412_275_275_29	NA
K02950	small subunit ribosomal protein S12	403_3947_3947_1	NA
K02950	small subunit ribosomal protein S12	409_3275_3275_1	NA
K03293	amino acid transporter, AAT family	403_976_976_25	NA
K03293	amino acid transporter, AAT family	403_2554_2554_1	NA
K03293	amino acid transporter, AAT family	403_1197_1197_1	NA
K03293	amino acid transporter, AAT family	409_9710_9710_1	NA
K03293	amino acid transporter, AAT family	403_1713_1713_1	NA
K03293	amino acid transporter, AAT family	403_8280_8280_2	NA
K03293	amino acid transporter, AAT family	405_1070_1070_13	NA
K03293	amino acid transporter, AAT family	405_673_673_12	NA
K03293	amino acid transporter, AAT family	405_279_279_7	NA
K03293	amino acid transporter, AAT family	405_1070_1070_14	NA
K03293	amino acid transporter, AAT family	403_2633_2633_2	NA
K03293	amino acid transporter, AAT family	416_1314_1314_1	NA
K03293	amino acid transporter, AAT family	403_8623_8623_15	NA
K03293	amino acid transporter, AAT family	403_2633_2633_3	NA
K03293	amino acid transporter, AAT family	403_2647_2647_1	NA
K03293	amino acid transporter, AAT family	409_13751_13751_2	NA
K03293	amino acid transporter, AAT family	409_8541_8541_1	NA
K03320	ammonium transporter, Amt family	403_957_957_6	NA
K03320	ammonium transporter, Amt family	403_8625_8625_9	NA
K03320	ammonium transporter, Amt family	405_260_260_5	NA
K03431	phosphoglucosamine mutase	403_1250_1250_3	NA
K03431	phosphoglucosamine mutase	403_3536_3536_2	NA
K03431	phosphoglucosamine mutase	403_3564_3564_1	NA
K03431	phosphoglucosamine mutase	405_250_250_11	NA
K03431	phosphoglucosamine mutase	403_3283_3283_2	NA
K03431	phosphoglucosamine mutase	412_380_380_21	NA
K03431	phosphoglucosamine mutase	403_3140_3140_1	NA
K03431	phosphoglucosamine mutase	409_141_141_14	NA
K03470	ribonuclease HII	403_1020_1020_1	NA
K03470	ribonuclease HII	403_2391_2391_4	NA
K03470	ribonuclease HII	405_279_279_24	NA
K03470	ribonuclease HII	403_5071_5071_1	NA
K03502	DNA polymerase V	405_234_234_12	NA
K03502	DNA polymerase V	416_639_639_1	NA
K03502	DNA polymerase V	403_8630_8630_5	NA
K03531	cell division protein FtsZ	403_8225_8225_3	NA
K03531	cell division protein FtsZ	403_8604_8604_5	NA
K03531	cell division protein FtsZ	409_581_581_166	NA
K03531	cell division protein FtsZ	405_375_375_12	NA
K03531	cell division protein FtsZ	403_4463_4463_1	NA
K03531	cell division protein FtsZ	409_10667_10667_1	NA
K03531	cell division protein FtsZ	409_8560_8560_1	NA
K03536	ribonuclease P protein component	403_954_954_22	NA
K03536	ribonuclease P protein component	403_8591_8591_2	NA
K03536	ribonuclease P protein component	405_211_211_7	NA
K03553	recombination protein RecA	403_1203_1203_4	NA
K03553	recombination protein RecA	403_2258_2258_1	NA
K03553	recombination protein RecA	403_2342_2342_3	NA
K03553	recombination protein RecA	405_240_240_19	NA
K03553	recombination protein RecA	403_2599_2599_9	NA

Table 17. (cont.)

K03553	recombination protein RecA	405_771_771_1	NA
K03553	recombination protein RecA	412_273_273_6	NA
K03553	recombination protein RecA	409_4643_4643_2	NA
K03553	recombination protein RecA	409_5950_5950_1	NA
K03553	recombination protein RecA	409_11841_11841_1	NA
K03584	DNA repair protein RecO (recombination protein O)	403_1063_1063_7	NA
K03584	DNA repair protein RecO (recombination protein O)	403_2262_2262_1	NA
K03584	DNA repair protein RecO (recombination protein O)	405_244_244_41	NA
K03584	DNA repair protein RecO (recombination protein O)	403_4774_4774_1	NA
K03589	cell division protein FtsQ	403_1239_1239_5	NA
K03589	cell division protein FtsQ	403_2363_2363_1	NA
K03589	cell division protein FtsQ	405_375_375_10	NA
K03647	protein involved in ribonucleotide reduction	403_989_989_2	NA
K03647	protein involved in ribonucleotide reduction	403_2457_2457_2	NA
K03647	protein involved in ribonucleotide reduction	403_8080_8080_3	NA
K03647	protein involved in ribonucleotide reduction	405_302_302_20	NA
K03647	protein involved in ribonucleotide reduction	403_2876_2876_2	NA
K03647	protein involved in ribonucleotide reduction	403_2852_2852_3	NA
K03978	GTP-binding protein	405_252_252_22	NA
K03978	GTP-binding protein	403_5363_5363_1	NA
K03978	GTP-binding protein	409_1972_1972_1	NA
K03978	GTP-binding protein	409_2804_2804_1	NA
K04066	primosomal protein N' (replication factor Y) (superfamily II helicase)	403_1178_1178_4	NA
K04066	primosomal protein N' (replication factor Y) (superfamily II helicase)	403_4515_4515_3	NA
K04066	primosomal protein N' (replication factor Y) (superfamily II helicase)	405_264_264_20	NA
K04066	primosomal protein N' (replication factor Y) (superfamily II helicase)	416_2340_2340_1	NA
K04066	primosomal protein N' (replication factor Y) (superfamily II helicase)	403_2637_2637_1	NA
K04066	primosomal protein N' (replication factor Y) (superfamily II helicase)	403_3196_3196_2	NA
K04066	primosomal protein N' (replication factor Y) (superfamily II helicase)	416_1379_1379_1	NA
K04066	primosomal protein N' (replication factor Y) (superfamily II helicase)	409_2389_2389_1	NA
K04094	glucose inhibited division protein Gid	405_279_279_21	NA
K04094	glucose inhibited division protein Gid	403_6286_6286_1	NA
K04094	glucose inhibited division protein Gid	409_6168_6168_1	NA
K05946	N-acetylglucosaminyldiphosphoundecaprenol	403_3127_3127_2	NA
K05946	N-acetylglucosaminyldiphosphoundecaprenol	405_247_247_6	NA
K06024	segregation and condensation protein B	405_242_242_7	NA
K06024	segregation and condensation protein B	403_2746_2746_3	NA
K06153	undecaprenyl-diphosphatase	416_2166_2166_1	NA
K06180	ribosomal large subunit pseudouridine synthase D	403_1241_1241_4	NA
K06180	ribosomal large subunit pseudouridine synthase D	403_8137_8137_2	NA
K06180	ribosomal large subunit pseudouridine synthase D	405_209_209_4	NA
K06180	ribosomal large subunit pseudouridine synthase D	405_356_356_6	NA
K06180	ribosomal large subunit pseudouridine synthase D	405_278_278_1	NA
K06180	ribosomal large subunit pseudouridine synthase D	403_3315_3315_3	NA
K06180	ribosomal large subunit pseudouridine synthase D	403_8593_8593_3	NA
K06180	ribosomal large subunit pseudouridine synthase D	403_8630_8630_16	NA
K06180	ribosomal large subunit pseudouridine synthase D	409_1434_1434_9	NA
K06180	ribosomal large subunit pseudouridine synthase D	409_2287_2287_4	NA
K06180	ribosomal large subunit pseudouridine synthase D	403_2088_2088_1	NA
K06180	ribosomal large subunit pseudouridine synthase D	409_1156_1156_1	NA
K06180	ribosomal large subunit pseudouridine synthase D	409_2372_2372_3	NA
K06213	magnesium transporter	403_5877_5877_1	NA
K06213	magnesium transporter	403_6549_6549_1	NA
K06213	magnesium transporter	409_4316_4316_2	NA
K07042	probable rRNA maturation factor	403_1165_1165_12	NA
K07042	probable rRNA maturation factor	403_2540_2540_9	NA
K07042	probable rRNA maturation factor	405_244_244_43	NA
K07738	transcriptional repressor NrdR	403_1100_1100_8	NA
K07738	transcriptional repressor NrdR	403_2770_2770_1	NA
K07738	transcriptional repressor NrdR	403_2908_2908_13	NA
K07738	transcriptional repressor NrdR	405_278_278_31	NA
K07738	transcriptional repressor NrdR	403_2636_2636_1	NA
K07738	transcriptional repressor NrdR	412_303_303_24	NA
K08483	phosphotransferase system, enzyme I, PtsI	403_1305_1305_5	NA
K08483	phosphotransferase system, enzyme I, PtsI	409_11932_11932_1	NA
K08483	phosphotransferase system, enzyme I, PtsI	405_209_209_12	NA

Table 17. (cont.)

K08483	phosphotransferase system, enzyme I, PtsI	405_209_209_12	NA
K08483	phosphotransferase system, enzyme I, PtsI	403_5076_5076_1	NA
K08483	phosphotransferase system, enzyme I, PtsI	403_8593_8593_11	NA
K08483	phosphotransferase system, enzyme I, PtsI	412_300_300_8	NA
K08483	phosphotransferase system, enzyme I, PtsI	403_438_438_3	NA
K09013	Fe-S cluster assembly ATP-binding protein	403_1025_1025_11	NA
K09013	Fe-S cluster assembly ATP-binding protein	403_8518_8518_3	NA
K09013	Fe-S cluster assembly ATP-binding protein	412_379_379_97	NA
K09013	Fe-S cluster assembly ATP-binding protein	416_2282_2282_1	NA
K09013	Fe-S cluster assembly ATP-binding protein	409_1730_1730_5	NA
K09013	Fe-S cluster assembly ATP-binding protein	409_2279_2279_4	NA
K09013	Fe-S cluster assembly ATP-binding protein	409_3666_3666_1	NA
K11031	thiol-activated cytolysin	403_2473_2473_1	NA
Note: NA represents that no information is provided.			

Table 18. Sample descriptions.

Sample ID	110	117	119	120	125	202	203	205	206
Race/Ethnicity	Caucasian	Caucasian	Caucasian	Caucasian	Caucasian	Caucasian	Caucasian	Caucasian	Caucasian
Age	34	43	31	33	23	30	35	33	29
Marital status	NA	Married	Married	Married	Married	Married	Married	Married	Married
Length of residence	more than 10 years	more than 10 years	more than 10 years	more than 10 years	1-2 years	2-5 years	1-2 years	2-5 years	2-5 years
Regular periods	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Length of menstrual cycle	29	30	28	29	27	28	30	30	32
Tampon use during periods	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of previous pregnancies	3	3	1	3	3	5	6	4	2
Number of children	1	0	0	2	1	2	4	2	1
Problems with the birth of any of your children	Failure to progress	NA	NA	NA	NA	PPROM, preterm	Preterm birth	PPROM, preterm	Preterm birth
Sexual orientation	Heterosexual	Heterosexual	Heterosexual	Heterosexual	Heterosexual	Heterosexual	Heterosexual	Heterosexual	Heterosexual
Sexual intercourse in the past 6 months	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of sexual partners in the past 6 months	1	1	1	1	1	1	1	1	1
New sexual partner in the past 6 months	No	No	No	No	No	No	No	No	No
New sexual partner in the past 2 months	No	No	No	No	No	No	No	No	No
Days since intercourse	7	0	7	4	2	3	6	7--14	60--180
Frequency of vaginal intercourse per month on average	Occasionally	Very Frequently	Occasionally	Very Frequently	Frequently	Occasionally	Occasionally	Rarely	Occasionally
Birth control methods	None	None	None	Birth control pill	Condom	None	None	None	Birth control pill
Douche	No	No	No	No	No	No	No	No	No
Diagnose with a vaginal bacterial infection	NA	NA	No	No	No	Yes	No	Yes	No

Note: NA represents no information is provided. Subject 201 did not turn in the questionnaire.

Table 19. Sequence statistics of metagenome vaginal samples.

Sample ID	Time point	Number of bases in raw metagenome	Absolute number of short reads in raw metagenome	Number of bases in quality-filtered metagenome	Absolute number of short reads in quality-filtered metagenome
110	8-12 wks	6,222,521,000	62,225,210	1,012,674,130	10,962,510
110	17-21 wks	3,419,244,400	34,192,444	548,610,851	5,967,898
110	26-30 wks	8,422,921,600	84,229,216	816,016,476	8,889,386
110	35-38 wks	6,153,614,600	61,536,146	1,449,913,587	15,725,802
110	Labor	6,521,101,000	65,211,010	808,972,001	8,784,180
110	6-wk post-partum	6,420,355,200	64,203,552	709,894,660	7,741,828
110	Total	37,159,757,800	371,597,578	5,346,081,705	58,071,604
117	8-12 wks	3,463,671,000	34,636,710	490,589,486	5,309,444
117	17-21 wks	2,769,467,800	27,694,678	802,458,979	8,697,176
117	26-30 wks	6,657,783,800	66,577,838	1,045,180,358	11,364,190
117	35-38 wks	6,033,391,800	60,333,918	1,010,236,534	10,977,160
117	Labor	6,532,720,200	65,327,202	919,926,088	10,016,428
117	6-wk post-partum	6,798,045,200	67,980,452	505,872,266	5,744,540
117	Total	32,255,079,800	322,550,798	4,774,263,711	52,108,938
119	8-12 wks	5,559,815,200	55,598,152	548,883,316	5,991,524
119	17-21 wks	6,532,671,400	65,326,714	808,238,308	8,733,978
119	26-30 wks	6,295,553,800	62,955,538	599,938,618	6,493,916
119	35-38 wks	5,682,729,000	56,827,290	693,469,800	7,329,562
119	6-wk post-partum	7,640,346,600	76,403,466	630,436,700	7,031,000
119	Total	31,711,116,000	317,111,160	3,280,966,742	35,579,980
120	8-12 wks	5,903,900,600	59,039,006	520,268,491	5,675,170
120	17-21 wks	6,788,770,400	67,887,704	650,151,325	7,248,860
120	26-30 wks	5,121,966,600	51,219,666	611,397,393	6,802,736
120	35-38 wks	5,662,146,400	56,621,464	707,210,925	7,699,112
120	6-wk post-partum	6,608,064,400	66,080,644	940,578,127	10,452,266
120	Total	30,084,848,400	300,848,484	3,429,606,261	37,878,144
125	8-12 wks	6,284,699,600	62,846,996	561,375,915	6,231,180
125	17-21 wks	5,341,225,600	53,412,256	489,532,965	5,445,938
125	26-30 wks	6,462,088,000	64,620,880	559,558,063	6,224,308
125	35-38 wks	5,670,923,200	56,709,232	501,003,132	5,446,978
125	6-wk post-partum	5,979,844,000	59,798,440	691,906,437	7,694,732
125	Total	29,738,780,400	297,387,804	2,803,376,512	31,043,136
201	8-12 wks	7,047,922,200	70,479,222	666,568,719	7,262,476
201	17-21 wks	6,189,886,000	61,898,860	677,123,851	7,370,600
201	26-30 wks	6,221,771,600	62,217,716	629,091,925	6,857,022
201	Labor	6,761,517,800	67,615,178	712,571,709	7,767,520
201	Total	26,221,097,600	262,210,976	2,685,356,204	29,257,618
202	8-12 wks	4,228,467,800	42,284,678	852,146,512	9,330,524
202	17-21 wks	6,497,271,000	64,972,710	866,593,791	9,641,248
202	26-30 wks	4,763,378,600	47,633,786	687,405,805	7,369,400
202	35-38 wks	5,922,797,600	59,227,976	751,314,007	8,033,618
202	6-wk post-partum	6,940,536,800	69,405,368	563,051,755	6,285,406
202	Total	28,352,451,800	283,524,518	3,720,511,870	40,660,196
203	8-12 wks	8,642,214,000	86,422,140	677,845,665	7,556,132
203	17-21 wks	6,585,760,800	65,857,608	739,270,944	7,760,692
203	26-30 wks	7,247,407,000	72,474,070	840,431,096	9,194,186
203	35-38 wks	6,312,381,600	63,123,816	674,927,563	7,351,998
203	Labor	8,013,985,600	80,139,856	968,043,839	10,734,354
203	6-wk post-partum	6,921,498,600	69,214,986	1,471,664,421	16,297,512
203	Total	43,723,247,600	437,232,476	5,372,183,528	58,894,874
205	8-12 wks	6,309,148,800	63,091,488	811,181,657	8,812,836
205	17-21 wks	8,844,356,800	88,443,568	1,283,524,268	13,953,034
205	26-30 wks	7,221,303,600	72,213,036	1,029,667,256	10,892,346
205	Labor	6,087,839,200	60,878,392	707,826,237	7,895,806
205	6-wk post-partum	8,321,430,600	83,214,306	1,426,745,772	15,501,576
205	Total	36,784,079,000	367,840,790	5,258,945,190	57,055,598
206	8-12 wks	7,745,705,400	77,457,054	641,737,529	7,156,076
206	17-21 wks	8,170,139,800	81,701,398	851,183,709	9,134,556
206	26-30 wks	8,335,937,400	83,359,374	958,816,544	10,158,590
206	35-38 wks	8,050,282,200	80,502,822	623,101,343	6,912,862
206	Labor	4,973,220,200	49,732,202	2,511,961,781	27,233,350
206	6-wk post-partum	7,366,342,400	73,663,424	733,724,833	8,169,590
206	Total	44,641,627,400	446,416,274	6,320,525,739	68,765,024

Table 19. (cont.)

Sample ID	Time point	Absolute number of short reads used in the recovery of the *Lactobacillus* population genomes	Number of short reads (normalized) used in the recovery of the *Lactobacillus* population genomes
110	8-12 wks	5,276,125	481,288
110	17-21 wks	1,491,031	249,842
110	26-30 wks	1,114,764	125,404
110	35-38 wks	10,911,413	693,854
110	Labor	2,547,933	290,059
110	6-wk post-partum	56,485	7,296
110	Total	21,397,751	1,847,743
117	8-12 wks	757,741	142,716
117	17-21 wks	6,277,029	721,732
117	26-30 wks	5,287,084	465,241
117	35-38 wks	5,523,602	503,190
117	Labor	3,865,682	385,934
117	6-wk post-partum	1,155	201
117	Total	21,712,293	2,219,014
119	8-12 wks	533,844	89,100
119	17-21 wks	1,430,332	163,766
119	26-30 wks	215,377	33,166
119	35-38 wks	287,509	39,226
119	6-wk post-partum	1,542	219
119	Total	2,468,604	325,477
120	8-12 wks	287	51
120	17-21 wks	1,659	229
120	26-30 wks	1,301	191
120	35-38 wks	2,113,355	274,493
120	6-wk post-partum	761	73
120	Total	2,117,363	275,037
125	8-12 wks	1,153,618	185,136
125	17-21 wks	748,341	137,413
125	26-30 wks	622,231	99,968
125	35-38 wks	580,446	106,563
125	6-wk post-partum	1,561	203
125	Total	3,106,197	529,283
201	8-12 wks	1,092,089	150,374
201	17-21 wks	824,061	111,804
201	26-30 wks	1,044,087	152,265
201	Labor	1,269,274	163,408
201	Total	4,229,511	577,851
202	8-12 wks	1,793,107	192,176
202	17-21 wks	738,746	76,623
202	26-30 wks	5,119,354	694,677
202	35-38 wks	486,455	60,552
202	6-wk post-partum	306	49
202	Total	8,137,968	1,024,078
203	8-12 wks	713	94
203	17-21 wks	401	52
203	26-30 wks	830,752	90,356
203	35-38 wks	698,779	95,046
203	Labor	253,340	23,601
203	6-wk post-partum	1,136	70
203	Total	1,785,121	209,219
205	8-12 wks	1,989,548	225,756
205	17-21 wks	6,534,130	468,295
205	26-30 wks	4,079,190	374,501
205	Labor	218,257	27,642
205	6-wk post-partum	9,045,875	583,546
205	Total	21,867,000	1,679,739
206	8-12 wks	713	100
206	17-21 wks	1,592	174
206	26-30 wks	1,907,960	187,817
206	35-38 wks	73,587	10,645
206	Labor	824,494	30,275
206	6-wk post-partum	1,208	148
206	Total	2,809,554	229,159

Table 19. (cont.)

Sample ID	Time point	Absolute number of short reads participated in the prediction of metabolism pathway in the *Lactobacillus* population genome	Number of short reads (normalized) participated in the prediction of metabolism pathway in the *Lactobacillus* population genome	Absolute number of bases aligned to assembly #1 and 2 (>= 500 bp)
110	8-12 wks	6,672	60,862	662,454,301
110	17-21 wks	3,254	54,525	185,311,020
110	26-30 wks	1,676	18,854	164,822,120
110	35-38 wks	5,563	35,375	1,060,057,727
110	Labor	2,242	25,523	270,937,143
110	6-wk post-partum	97	1,253	41,480,454
110	Total	19,504	196,392	2,385,062,765
117	8-12 wks	1,192	22,451	89,430,036
117	17-21 wks	6,067	69,758	593,656,074
117	26-30 wks	3,923	34,521	513,101,161
117	35-38 wks	4,155	37,851	533,456,698
117	Labor	3,222	32,167	386,358,559
117	6-wk post-partum	0	0	66,556,987
117	Total	18,559	196,748	2,182,559,515
119	8-12 wks	712	11,883	71,292,101
119	17-21 wks	1,281	14,667	163,230,890
119	26-30 wks	263	4,050	44,797,528
119	35-38 wks	302	4,120	54,964,775
119	6-wk post-partum	0	0	31,138,980
119	Total	2,558	34,721	365,424,274
120	8-12 wks	0	0	31,217,350
120	17-21 wks	0	0	139,004,233
120	26-30 wks	0	0	178,606,156
120	35-38 wks	2,406	31,250	229,096,007
120	6-wk post-partum	0	0	523,555,998
120	Total	2,406	31,250	1,101,479,744
125	8-12 wks	2,122	34,055	158,360,061
125	17-21 wks	1,569	28,810	94,199,754
125	26-30 wks	1,145	18,396	85,599,898
125	35-38 wks	1,231	22,600	90,320,915
125	6-wk post-partum	1	13	320,181,576
125	Total	6,068	103,873	748,662,204
201	8-12 wks	2,371	32,647	173,024,382
201	17-21 wks	1,723	23,377	132,615,712
201	26-30 wks	2,365	34,490	160,300,236
201	Labor	2,510	32,314	188,866,362
201	Total	8,969	122,828	654,806,692
202	8-12 wks	2,080	22,292	260,673,767
202	17-21 wks	839	8,702	241,340,670
202	26-30 wks	179	2,429	528,676,059
202	35-38 wks	664	8,265	213,756,723
202	6-wk post-partum	0	0	33,654,931
202	Total	3,762	41,689	1,278,102,150
203	8-12 wks	0	0	98,601,592
203	17-21 wks	0	0	108,616,190
203	26-30 wks	873	9,495	136,155,283
203	35-38 wks	938	12,758	112,201,721
203	Labor	230	2,143	125,966,244
203	6-wk post-partum	0	0	62,392,928
203	Total	2,041	24,396	643,933,958
205	8-12 wks	1,748	19,835	213,816,916
205	17-21 wks	3,593	25,751	638,504,441
205	26-30 wks	2,826	25,945	425,544,582
205	Labor	221	2,799	52,763,657
205	6-wk post-partum	4,441	28,649	874,736,344
205	Total	12,829	102,978	2,205,365,940
206	8-12 wks	0	0	175,667,171
206	17-21 wks	0	0	163,946,678
206	26-30 wks	2,188	21,538	297,809,016
206	35-38 wks	126	1,823	96,626,772
206	Labor	369	1,355	2,225,820,070
206	6-wk post-partum	0	0	93,301,014
206	Total	2,683	24,716	3,053,170,721

Table 20. Summary of the percentage of bases covered for each HMP urogenital tract reference genome in each sample.

	110	117	119	120	125	201	202	203	205	206
Acinetobacter_baumannii_ATCC_19606_=_CIP_70.34	0.01%	0.00%	0.00%	0.08%	0.01%	0.00%	0.01%	0.01%	0.01%	0.03%
Actinobaculum_massiliae_ACS-171-V-Col2	0.01%	0.01%	0.04%	0.03%	0.01%	0.00%	0.02%	0.48%	4.10%	0.01%
Actinomyces_coleocanis_DSM_15436	0.00%	0.02%	0.01%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Actinomyces_neuii_BVS029A5	0.15%	0.00%	0.16%	0.46%	0.01%	0.04%	0.01%	0.59%	1.03%	0.17%
Actinomyces_turicensis_ACS-279-V-Col4	0.92%	0.00%	0.02%	0.35%	0.13%	0.01%	19.67%	5.56%	0.49%	0.55%
Actinomyces_urogenitalis_DSM_15434	0.00%	0.00%	0.00%	0.02%	0.04%	0.00%	0.06%	0.03%	0.29%	0.01%
Aerococcus_urinae_ACS-120-V-Col10a	0.01%	0.00%	0.00%	0.01%	0.00%	0.00%	0.02%	0.07%	0.12%	0.00%
Aerococcus_viridans_ATCC_11563	0.00%	0.01%	0.00%	0.03%	0.00%	0.00%	0.00%	0.01%	0.01%	0.01%
Anaerococcus_hydrogenalis_ACS-025-V-Sch4	1.49%	0.96%	0.25%	7.75%	2.61%	0.04%	1.79%	15.49%	1.22%	0.78%
Anaerococcus_lactolyticus_ATCC_51172	1.01%	0.05%	0.71%	0.37%	12.02%	0.02%	5.08%	49.06%	0.30%	0.38%
Anaerococcus_prevotii_ACS-065-V-Col13	0.37%	0.31%	0.66%	1.06%	0.35%	0.07%	1.31%	2.33%	0.02%	0.60%
Anaerococcus_tetradius_ATCC_35098	0.08%	0.02%	2.34%	0.23%	4.86%	0.01%	0.64%	0.39%	0.06%	0.09%
Atopobium_vaginae_DSM_15829	0.01%	0.02%	0.04%	0.08%	91.93%	10.64%	4.69%	0.00%	0.05%	0.12%
Atopobium_vaginae_PB189-T1-4	0.00%	0.00%	0.00%	0.00%	1.04%	0.09%	0.05%	0.09%	0.00%	0.08%
Bifidobacterium_breve_ACS-071-V-Sch8b	0.00%	20.26%	15.00%	0.07%	5.59%	0.01%	0.08%	5.43%	0.11%	0.55%
Bifidobacterium_dentium_ATCC_27679	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Bifidobacterium_dentium_JCVIHMP022	0.00%	0.01%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Brevibacterium_mcbrellneri_ATCC_49030	0.00%	0.02%	0.01%	0.28%	0.00%	0.00%	0.00%	0.04%	0.16%	0.01%
Chryseobacterium_gleum_ATCC_35910	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	0.01%
Clostridiales_genomosp._BVAB3_str._UPII9-5	0.13%	0.04%	0.09%	0.18%	0.73%	0.08%	0.32%	0.39%	0.03%	0.02%
Corynebacterium_aurimucosum_ATCC_700975	0.00%	0.01%	0.03%	1.27%	0.21%	0.01%	0.02%	0.41%	1.11%	0.04%
Corynebacterium_genitalium_ATCC_33030	0.00%	0.34%	0.00%	0.02%	0.00%	0.00%	0.01%	0.12%	0.44%	0.01%
Corynebacterium_glucuronolyticum_ATCC_51866	0.00%	0.01%	0.00%	0.38%	0.00%	0.01%	0.02%	0.10%	0.06%	0.01%
Corynebacterium_glucuronolyticum_ATCC_51867	0.00%	0.00%	0.00%	0.25%	0.00%	0.00%	0.01%	0.09%	0.01%	0.00%
Corynebacterium_jeikeium_ATCC_43734	0.00%	0.52%	0.09%	0.94%	0.01%	0.00%	0.01%	2.76%	1.08%	0.37%
Corynebacterium_lipophiloflavum_DSM_44291	0.00%	0.01%	0.01%	0.01%	0.00%	0.00%	0.01%	0.01%	0.07%	0.00%
Corynebacterium_pseudogenitalium_ATCC_33035	0.25%	0.92%	0.20%	3.14%	1.50%	0.08%	0.58%	5.41%	3.61%	1.83%
Corynebacterium_striatum_ATCC_6940	0.00%	0.00%	0.00%	0.27%	0.03%	0.00%	0.04%	0.04%	0.45%	0.30%
Dialister_micraerophilus_DSM_19965	0.09%	0.49%	1.88%	0.71%	1.45%	0.00%	0.78%	0.95%	0.05%	0.20%
Dialister_microaerophilus_UPII_345-E	0.00%	0.04%	0.23%	0.20%	0.06%	0.07%	0.25%	0.07%	0.04%	0.00%
Enterococcus_durans_FB129-CNAB-4	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%
Enterococcus_faecalis_ATCC_29200	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_ERV81	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_ERV85	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_ERV93	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_HH22	0.00%	0.03%	0.00%	0.01%	0.00%	0.00%	0.00%	0.03%	0.01%	0.01%
Enterococcus_faecalis_TX0312	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_TX0635	0.00%	0.01%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecalis_TX0855	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Enterococcus_faecium_E422	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%
Enterococcus_faecium_P1139	0.00%	0.00%	0.21%	0.00%	0.00%	0.01%	0.03%	0.00%	0.01%	0.00%
Enterococcus_faecium_S447	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Eremococcus_coleocola_ACS-139-V-Col8	0.03%	0.09%	0.00%	0.22%	0.01%	0.03%	0.03%	0.19%	0.02%	0.01%
Escherichia_coli_83972	4.46%	2.74%	3.19%	67.29%	19.74%	3.40%	2.59%	4.73%	6.59%	5.73%
Finegoldia_magna_ACS-171-V-Col3	0.34%	0.14%	0.28%	1.77%	0.52%	0.03%	0.69%	2.99%	0.41%	0.16%
Finegoldia_magna_ATCC_53516	0.33%	1.47%	0.03%	2.01%	0.32%	0.09%	0.42%	1.14%	0.68%	1.08%
Finegoldia_magna_BVS033A4	0.30%	0.02%	0.30%	2.02%	0.55%	0.03%	0.75%	2.95%	0.31%	0.22%
Finegoldia_magna_SY403409CC001050417	0.37%	0.08%	0.32%	1.73%	0.45%	0.04%	0.94%	2.59%	0.20%	0.17%
Fusobacterium_nucleatum_subsp._nucleatum_ATCC_23726	0.02%	5.16%	0.10%	0.19%	0.02%	0.00%	0.31%	0.08%	0.58%	0.03%
Gardnerella_vaginalis_315-A	0.05%	0.10%	0.09%	18.75%	8.91%	0.03%	0.30%	18.50%	0.59%	1.89%
Gardnerella_vaginalis_409-05	18.59%	0.03%	2.61%	19.03%	4.95%	13.00%	94.37%	4.56%	23.29%	9.25%
Gardnerella_vaginalis_ATCC_14019	0.03%	0.16%	0.08%	19.47%	8.40%	0.02%	0.34%	17.12%	0.55%	1.31%
Gardnerella_vaginalis_HMP9231	0.23%	0.14%	0.14%	20.73%	8.15%	0.41%	2.92%	17.15%	0.79%	1.46%
Haemophilus_parainfluenzae_HK262	0.00%	0.00%	0.00%	0.08%	0.00%	0.00%	0.13%	23.41%	0.08%	0.06%
Klebsiella_oxytoca_10-5245	0.07%	0.03%	0.04%	0.21%	0.04%	0.04%	0.03%	0.05%	0.06%	0.08%
Klebsiella_oxytoca_10-5250	0.00%	0.00%	0.00%	0.12%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%
Lactobacillus_coleohominis_101-4-CHN	5.54%	0.04%	0.00%	0.40%	1.79%	0.06%	0.00%	0.47%	0.79%	0.00%
Lactobacillus_crispatus_125-2-CHN	1.77%	4.76%	0.53%	3.24%	0.14%	0.00%	0.02%	0.11%	1.65%	0.03%
Lactobacillus_crispatus_214-1	1.93%	2.65%	0.87%	0.91%	0.02%	0.00%	0.02%	0.05%	2.00%	0.10%
Lactobacillus_crispatus_CTV-05	5.48%	4.28%	5.94%	2.14%	0.04%	0.02%	0.02%	0.09%	7.80%	0.83%
Lactobacillus_crispatus_FB049-03	3.71%	2.89%	0.88%	1.62%	0.02%	0.01%	0.02%	0.02%	4.24%	0.51%
Lactobacillus_crispatus_FB077-07	4.34%	1.77%	4.45%	0.73%	0.01%	0.04%	0.16%	0.06%	6.22%	0.73%
Lactobacillus_crispatus_JV-V01	0.60%	1.52%	0.14%	1.34%	0.08%	0.00%	0.00%	0.02%	0.81%	0.01%
Lactobacillus_crispatus_MV-1A-US	0.75%	1.19%	0.48%	0.92%	0.00%	0.00%	0.01%	0.00%	1.07%	0.04%

Table 20. (cont.)

Lactobacillus_crispatus_MV-3A-US	1.25%	2.63%	0.63%	2.03%	0.06%	0.00%	0.06%	0.01%	1.73%	0.03%
Lactobacillus_crispatus_SJ-3C-US	1.84%	2.53%	1.32%	1.36%	0.05%	0.00%	0.01%	0.11%	2.19%	0.09%
Lactobacillus_delbrueckii_subsp._bulgaricus_PB2003/044-T3-4	0.02%	0.06%	0.01%	0.00%	0.00%	0.00%	0.00%	0.02%	0.01%	0.04%
Lactobacillus_fermentum_28-3-CHN	0.02%	0.03%	0.01%	5.81%	0.15%	0.02%	0.02%	49.24%	0.08%	0.08%
Lactobacillus_gasseri_202-4	0.86%	0.06%	0.14%	0.65%	19.75%	2.17%	29.74%	0.26%	5.84%	0.87%
Lactobacillus_gasseri_224-1	0.09%	0.07%	0.00%	1.00%	0.31%	0.06%	3.47%	0.02%	0.04%	1.61%
Lactobacillus_gasseri_JV-V03	0.05%	0.03%	0.04%	0.37%	0.30%	0.04%	1.85%	0.61%	0.15%	0.66%
Lactobacillus_gasseri_MV-22	0.40%	0.15%	0.07%	1.91%	0.11%	0.04%	2.00%	0.00%	0.43%	2.89%
Lactobacillus_gasseri_SJ-9E-US	0.05%	0.09%	0.05%	2.70%	0.17%	0.00%	0.16%	0.01%	0.19%	3.08%
Lactobacillus_gasseri_SV-16A-US	0.06%	0.01%	0.01%	0.09%	0.88%	0.16%	1.70%	0.00%	0.19%	0.12%
Lactobacillus_iners_ATCC_55195	5.31%	0.01%	0.00%	0.03%	4.90%	5.05%	0.07%	0.03%	0.05%	0.05%
Lactobacillus_iners_DSM_13335	4.13%	0.00%	0.00%	0.03%	4.49%	4.94%	0.06%	0.00%	0.05%	0.09%
Lactobacillus_iners_LactinV_01V1-a	3.34%	0.01%	0.04%	0.24%	4.73%	6.30%	0.27%	0.01%	0.01%	0.04%
Lactobacillus_iners_LactinV_03V1-b	5.39%	0.03%	0.12%	0.01%	3.12%	4.18%	0.10%	0.00%	0.01%	0.07%
Lactobacillus_iners_LactinV_09V1-c	4.12%	0.01%	0.01%	0.03%	2.67%	4.21%	0.07%	0.05%	0.04%	0.05%
Lactobacillus_iners_LactinV_11V1-d	3.34%	0.01%	0.03%	0.06%	2.88%	4.24%	0.08%	0.00%	0.03%	0.09%
Lactobacillus_iners_LEAF_2052A-d	5.08%	0.00%	0.03%	0.07%	3.31%	4.32%	0.05%	0.00%	0.04%	0.00%
Lactobacillus_iners_LEAF_2053A-b	4.68%	0.00%	0.02%	0.07%	3.88%	5.00%	0.04%	0.01%	0.04%	0.06%
Lactobacillus_iners_LEAF_2062A-h1	4.86%	0.01%	0.03%	0.00%	2.84%	4.20%	0.04%	0.14%	0.00%	0.01%
Lactobacillus_iners_LEAF_3008A-a	3.60%	0.00%	0.00%	0.00%	2.40%	3.55%	0.07%	0.01%	0.03%	0.00%
Lactobacillus_iners_SPIN_1401G	5.08%	0.01%	0.01%	0.07%	3.71%	3.55%	0.05%	0.01%	0.05%	0.04%
Lactobacillus_iners_SPIN_2503V10-D	5.62%	0.00%	0.01%	0.06%	2.44%	4.89%	0.06%	0.03%	0.04%	0.09%
Lactobacillus_iners_UPII_143-D	2.84%	0.00%	0.02%	0.03%	1.79%	2.96%	0.12%	0.02%	0.01%	0.01%
Lactobacillus_iners_UPII_60-B	6.18%	0.04%	0.03%	0.13%	5.80%	5.53%	0.43%	0.04%	0.06%	0.10%
Lactobacillus_jensenii_1153	0.00%	0.23%	0.00%	0.01%	0.00%	0.00%	0.03%	0.00%	0.02%	0.00%
Lactobacillus_jensenii_115-3-CHN	0.05%	0.19%	0.00%	0.06%	4.06%	0.00%	0.52%	0.00%	0.00%	0.02%
Lactobacillus_jensenii_269-3	0.00%	0.04%	0.00%	0.02%	0.93%	0.01%	0.08%	0.00%	0.01%	0.00%
Lactobacillus_jensenii_27-2-CHN	0.00%	0.13%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.03%	0.00%
Lactobacillus_jensenii_JV-V16	0.06%	1.75%	0.07%	0.06%	0.13%	0.00%	0.11%	0.00%	0.05%	0.00%
Lactobacillus_jensenii_SJ-7A-US	0.22%	0.64%	0.02%	0.05%	2.22%	0.06%	0.02%	0.00%	0.06%	0.00%
Lactobacillus_johnsonii_ATCC_33200	0.04%	0.11%	0.01%	0.12%	0.06%	0.00%	0.25%	0.01%	0.03%	0.15%
Lactobacillus_oris_PB013-T2-3	0.01%	0.01%	0.00%	1.68%	0.11%	0.25%	0.00%	1.00%	0.10%	0.13%
Lactobacillus_salivarius_ACS-116-V-Col5a	0.00%	0.00%	0.04%	0.03%	0.00%	0.00%	0.02%	0.05%	0.01%	0.03%
Lactobacillus_vaginalis_ATCC_49540	0.11%	48.98%	0.10%	82.81%	24.70%	0.04%	0.00%	0.20%	33.63%	0.16%
Megasphaera_genomosp._type_1_str._28L	0.02%	0.01%	0.09%	0.10%	0.32%	0.01%	0.13%	0.46%	0.00%	0.05%
Megasphaera_sp._UPII_135-E	0.03%	0.02%	0.04%	0.12%	0.18%	0.01%	0.19%	0.65%	0.02%	0.07%
Megasphaera_sp._UPII_199-6	0.02%	0.00%	0.96%	0.55%	0.08%	0.02%	0.04%	0.24%	0.00%	0.02%
Mobiluncus_curtisii_ATCC_43063	0.00%	0.00%	0.00%	0.07%	0.06%	0.00%	0.08%	0.03%	0.02%	0.01%
Mobiluncus_curtisii_ATCC_51333	0.00%	0.00%	0.00%	0.66%	0.03%	0.00%	0.15%	0.04%	0.03%	0.07%
Mobiluncus_curtisii_subsp._curtisii_ATCC_35241	0.00%	0.00%	0.00%	0.01%	0.01%	0.00%	0.02%	0.00%	0.00%	0.03%
Mobiluncus_curtisii_subsp._holmesii_ATCC_35242	0.00%	0.00%	0.00%	0.04%	0.03%	0.00%	0.10%	0.02%	0.02%	0.19%
Mobiluncus_mulieris_28-1	0.00%	0.00%	0.00%	0.01%	0.01%	0.00%	0.00%	0.01%	0.00%	0.01%
Mobiluncus_mulieris_ATCC_35239	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Mobiluncus_mulieris_ATCC_35243	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.01%	0.00%
Mobiluncus_mulieris_FB024-16	0.00%	0.01%	0.00%	0.01%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%
Mycobacterium_parascrofulaceum_ATCC_BAA-614	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Myroides_odoratimimus_CCUG_12901	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Myroides_odoratimimus_CCUG_3837	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Pasteurella_bettyae_CCUG_2042	0.00%	0.00%	0.01%	0.02%	0.07%	0.00%	0.03%	0.14%	0.00%	0.01%
Peptoniphilus_duerdenii_ATCC_BAA-1640	0.13%	0.12%	0.81%	2.48%	1.26%	0.07%	6.09%	14.20%	0.73%	0.35%
Peptoniphilus_harei_ACS-146-V-Sch2b	2.35%	1.02%	1.33%	7.59%	2.23%	0.12%	11.15%	19.48%	5.63%	0.35%
Peptoniphilus_lacrimalis_315-B	0.09%	0.08%	0.88%	0.53%	2.45%	0.04%	3.31%	13.76%	1.08%	0.30%
Porphyromonas_asaccharolytica_PR426713P-I	0.01%	0.04%	0.34%	0.05%	0.33%	0.00%	3.85%	1.31%	0.00%	0.39%
Porphyromonas_uenonis_60-3	0.03%	0.20%	0.06%	0.66%	1.91%	0.00%	17.33%	0.61%	0.10%	1.90%
Prevotella_amnii_CRIS_21A-A	0.02%	0.05%	0.06%	0.05%	0.34%	0.01%	0.25%	0.28%	0.02%	1.91%
Prevotella_bivia_JCVIHMP010	0.39%	0.11%	3.80%	1.26%	15.06%	0.17%	0.61%	0.36%	0.71%	85.42%
Prevotella_buccalis_ATCC_35310	0.44%	1.00%	0.41%	1.75%	5.48%	0.04%	12.07%	2.56%	0.26%	4.99%
Prevotella_denticola_CRIS_18C-A	0.02%	0.02%	0.02%	0.01%	0.29%	0.01%	0.10%	0.19%	0.03%	0.70%
Prevotella_disiens_FB035-09AN	0.41%	0.49%	1.80%	1.13%	1.90%	0.10%	1.10%	2.47%	1.13%	2.94%
Prevotella_oralis_ATCC_33269	0.00%	0.00%	0.00%	0.00%	0.08%	0.00%	0.00%	0.00%	0.00%	0.03%
Prevotella_timonensis_CRIS_5C-B1	3.75%	8.68%	4.04%	20.36%	49.66%	0.18%	6.42%	22.31%	2.60%	23.47%
Propionibacterium_sp._409-HC1	0.24%	0.07%	0.06%	0.13%	0.01%	0.03%	0.07%	0.08%	0.41%	1.28%
Propionibacterium_sp._434-HC2	0.07%	0.10%	0.02%	0.07%	0.01%	0.02%	0.02%	0.04%	0.35%	0.64%
Proteus_mirabilis_ATCC_29906	0.00%	0.00%	0.00%	0.04%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Roseomonas_cervicalis_ATCC_49957	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Sphingobacterium_spiritivorum_ATCC_33300	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%
Sphingobacterium_spiritivorum_ATCC_33861	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Staphylococcus_aureus_subsp._aureus_MN8	0.00%	1.71%	0.00%	0.03%	0.01%	0.00%	0.01%	0.02%	0.01%	0.00%
Staphylococcus_epidermidis_BVS058A4	0.15%	0.75%	0.06%	9.33%	0.67%	0.11%	0.70%	4.95%	3.77%	0.44%

Table 20. (cont.)

Staphylococcus_lugdunensis_ACS-027-V-Sch2	0.00%	0.01%	0.01%	0.23%	0.01%	0.02%	0.00%	0.01%	0.03%	0.00%
Staphylococcus_simulans_ACS-120-V-Sch1	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.01%	0.03%	0.00%
Streptococcus_bovis_ATCC_700338	0.03%	0.02%	0.73%	0.38%	0.01%	0.01%	0.11%	0.84%	0.02%	0.40%
Streptococcus_pseudoporcinus_SPIN_20026	0.11%	0.05%	0.43%	0.38%	0.13%	0.00%	0.52%	1.23%	0.04%	0.34%
Streptococcus_urinalis_FB127-CNA-2	0.17%	0.01%	1.53%	0.68%	0.02%	0.01%	0.17%	2.04%	0.04%	0.20%
Treponema_phagedenis_F0421	0.00%	0.00%	0.00%	0.01%	0.01%	0.00%	0.00%	0.01%	0.01%	0.00%
Veillonella_atypica_ACS-049-V-Sch6	0.00%	0.05%	0.00%	0.21%	0.00%	0.00%	0.00%	0.16%	0.00%	0.00%
Veillonella_atypica_ACS-134-V-Col7a	0.01%	0.03%	0.00%	0.23%	0.01%	0.00%	0.00%	0.27%	0.00%	0.00%
Veillonella_parvula_ACS-068-V-Sch12	0.04%	0.73%	0.00%	0.50%	0.01%	0.01%	0.03%	22.27%	0.02%	0.01%
Veillonella_ratti_ACS-216-V-Col6b	0.01%	0.01%	0.01%	0.12%	0.00%	0.00%	0.10%	0.05%	0.00%	0.02%
Note: only reads (mapping quality > 20) were considered as correct alignment.										

Table 21. Assembly statistics of metagenome vaginal samples.

Sample ID	110	117	119	120	125	201	202	203	205	206
Number of contigs in assembly #1	383	307	136	1,895	704	267	372	444	246	767
Total size of contigs in assembly #1	342,412	311,577	189,805	4,891,172	1,704,923	237,337	1,982,114	405,458	236,882	3,306,810
Longest contig length in assembly #1 (bp)	9,730	6,599	41,629	42,826	54,768	6,876	199,022	10,517	6,900	48,481
N50 contig length in assembly #1 (bp)	898	1,000	1,660	6,391	6,790	873	59,497	921	942	9,956
Number of contigs in assembly #2	2,090	3,079	2,573	3,840	3,905	1,306	1,509	4,137	2,174	2,959
Total size of contigs in assembly #2 (bp)	5,560,865	6,612,830	4,464,242	10,122,789	7,001,139	2,577,370	5,118,034	7,562,140	4,309,859	9,125,125
Longest contig length in assembly #2 (bp)	152,421	60,015	137,548	535,652	86,365	192,112	343,435	120,010	104,053	375,197
N50 contig length in assembly #2 (bp)	13,320	6,377	4,859	10,027	2,908	21,712	40,086	5,341	5,341	46,037
Number of contigs in assembly #1 and #2	2,473	3,386	2,709	5,735	4,609	1,573	1,881	4,581	2,420	3,726
Total size of contigs in assembly #1 and #2 (bp)	5,903,277	6,924,407	4,654,047	15,013,961	8,706,062	2,814,707	7,100,148	7,967,598	4,546,741	12,431,935
Longest contig length in assembly #1 and #2 (bp)	152,421	60,015	137,548	535,652	86,365	192,112	343,435	120,010	104,053	375,197
N50 contig length in assembly #1 and #2 (bp)	11,653	5,803	4,609	7,815	3,836	6,490	41,768	3,874	4,438	17,469
Number of population genome bins	9	8	7	12	12	23	12	8	2	18

Note: Assemblies were composed of contigs that are >= 500 bp. Contig N50 is defined as the contig length such that using equal or longer contigs produces 50% the bases of the assembly.

Table 22. Summary statistics for high-quality population genomes.

Subject ID	Bin ID	Genome name	Number of contigs	Total size of contigs (bp)	Longest contig length (bp)	N50 contig length (bp)	Completeness (%)	Contamination (%)
110	1	*Lactobacillus*	313	2,334,097	70,760	19,920	99.03	1.44
110	25	*Lactobacillus*	180	1,254,532	137,752	73,573	94.62	8.2
117	4	*Lactobacillus*	297	2,172,438	60,015	19,888	99.03	4.19
119	4	*Lactobacillus*	307	2,435,439	137,548	18,797	99.03	2.85
120	4	*Lactobacillus*	316	2,217,935	83,781	17,332	99.03	0.95
125	8	*Lactobacillus*	162	1,610,216	67,625	18,776	97.06	2.81
125	22	*Lactobacillus*	157	1,310,644	86,365	49,793	96.09	8.16
201	40	*Lactobacillus*	151	1,141,992	192,112	93,741	87.99	7.99
202	31	*Lactobacillus*	67	1,691,880	155,626	45,230	89.15	0
203	4	*Lactobacillus*	107	2,186,361	120,010	30,962	97.93	0.65
205	1	*Lactobacillus*	414	2,525,300	104,053	15,571	98.17	3.69
206	16	*Lactobacillus*	29	1,760,305	375,197	265,736	99.22	0

Table 23. Taxonomic abundance determined from the metagenomic short reads in women with and without a previous preterm birth history (%).

Bacterial species	110_1	110_2	110_3	110_4	110_5	110_6	117_1	117_2	117_3	117_4	117_5
[Bacteroides]_pectinophilus	0	0	0	0	0	0	0	0	0	0	0
[Clostridium]_bartlettii	0	0	0	0	0	0	0	0	0	0	0
[Clostridium]_sticklandii	0	0	0	0	0	0	0	0	0	0	0
[Ruminococcus]_gnavus	0	0	0	0	0	0	0	0	0	0	0
[Ruminococcus]_obeum	0	0	0	0	0	0	0	0	0	0	0
[Ruminococcus]_torques	0	0	0	0	0	0	0	0	0	0	0
Acetobacterium_woodii	0	0	0	0	0	0	0	0	0	0	0
Acholeplasma_laidlawii	0	0	0	0	0	0	0	0	0	0	0
Acidimicrobium_ferrooxidans	0	0	0	0	0	0	0	0	0	0	0
Acidothermus_cellulolyticus	0	0	0	0	0	0	0	0	0	0	0
Aciduliprofundum_boonei	0	0	0	0	0	0	0	0	0	0	0
Acinetobacter_baumannii	0	0	0	0	0	0	0	0	0	0	0
Acinetobacter_junii	0	0	0	0	0	0	0	0	0	0	0
Acinetobacter_lwoffii	0	0	0	0	0	0	0	0	0	0	0
Actinobacillus_succinogenes	0	0	0	0	0	0	2.00E-02	0	1.00E-02	1.00E-02	0
Actinomyces_coleocanis	0	0	0	0	0	0	0	0	0	0	0
Actinomyces_urogenitalis	0	0	0	0	0	0	0	0	0	0	0
Aggregatibacter_aphrophilus	0	0	0	0	0	0	0	0	0	0	0
Alistipes_finegoldii	0	0	0	0	0	0	0	0	0	0	0
Alistipes_shahii	0	0	0	0	0	0	0	0	0	0	0
Alkaliphilus_metalliredigens	0	0	0	0	0	0	0	0	0	0	0
Allochromatium_vinosum	0	0	0	0	0	0	0	0	0	0	0
Alteromonas_macleodii	0	0	0	0	0	0	0	0	0	0	0
Alteromonas_sp._sn2	0	0	0	0	0	0	0	0	0	0	0
Amycolatopsis_mediterranei	0	0	0	0	0	0	0	0	0	0	0
Amycolicicoccus_subflavus	0	0	0	0	0	0	0	0	0	0	0
Anaerococcus_hydrogenalis	0	0	0	0	0	0.87	0	0	0	0	0
Anaerococcus_lactolyticus	0	0	0	0	1.00E-02	2.7	0	0	0	0	0
Anaerococcus_prevotii	0	0	1.00E-02	0	0	0	0	0	0	0	0
Anaerococcus_tetradius	0	0	0	0	0	0	0	0	0	0	0
Anaerococcus_vaginalis	0	0	0	0	0	1.06	2.00E-02	0	0	0	1.00E-02
Anaerotruncus_colihominis	0	0	0	0	0	0	0	0	0	0	0
Arcobacter_nitrofigilis	0	0	0	0	0	0	0	0	0	0	0
Atopobium_parvulum	0	0	0	0	0	0	0	0	0	0	0
Atopobium_vaginae	0	0	0	0	0	0	0	0	0	0	0
Azospira_oryzae	0	0	0	0	0	0	0	0	0	0	0
Bacillus_cereus	0	0	0	0	0	0	0	0	0	0	0
Bacillus_clausii	0	0	0	0	0	0	0	0	0	0	0
Bacillus_pseudofirmus	0	0	0	0	0	0	0	0	0	0	0
Bacillus_thuringiensis	0	0	0	0	0	0	0	0	0	0	0
Bacteriovorax_marinus	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_caccae	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_cellulosilyticus	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_dorei	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_fragilis	0	0	0	0	0	0.1	0	0	0	0	0
Bacteroides_ovatus	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_stercoris	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_thetaiotaomicron	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_uniformis	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_vulgatus	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_adolescentis	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_angulatum	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_animalis	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_bifidum	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_breve	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_catenulatum	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_dentium	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_gallicum	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_longum	0	0	0	0	0	0	0	0	0	0	1.00E-02
Bifidobacterium_pseudocatenulatum	0	0	0	0	0	0	0	0	0	0	0
Bilophila_wadsworthia	0	0	0	0	0	0	0	0	0	0	0
Blautia_hydrogenotrophica	0	0	0	0	0	0	0	0	0	0	0
Brevibacterium_mcbrellneri	0	0	0	0	0	0	0	0	0	0	1.00E-02

Table 23. (cont.)

Buchnera_aphidicola	0	0	0	0	0	0	0	0	0	0	0
Burkholderia_glumae	0	0	0	0	0	0	0	0	0	0	0
Burkholderia_phytofirmans	0	0	0	0	0	0	0	0	0	0	0
Burkholderia_pseudomallei	0	0	0	0	0	0	0	2.00E-02	1.00E-02	1.00E-02	1.00E-02
Butyrivibrio_crossotus	0	0	0	0	0	0	0	0	0	0	0
Butyrivibrio_fibrisolvens	0	0	0	0	0	0	0	0	0	0	0
Butyrivibrio_proteoclasticus	0	0	0	0	0	0	0	0	0	0	0
Caldicellulosiruptor_bescii	0	0	0	0	0	0	0	0	0	0	0
Campylobacter_fetus	0	0	0	0	0	0	0	0	0	0	0
Campylobacter_gracilis	0	0	0	0	0	0	0	0	0	0	0
Campylobacter_hominis	0	0	0	0	0	0	0	0	0	0	0
Campylobacter_lari	0	0	0	0	0	0	0	0	0	0	0
Campylobacter_upsaliensis	0	0	0	0	0	0.1	0	0	0	0	0
Candidatus_accumulibacter_phosphatis	0	0	0	0	0	0	0	0	0	0	0
Candidatus_azobacteroides_pseudotrichonymphae	0	0	0	0	0	0	0	0	0	0	0
Candidatus_baumannia_cicadellinicola	0	0	0	2.00E-02	1.00E-02	0	4.00E-02	2.00E-02	2.00E-02	2.00E-02	1.00E-02
Candidatus_carsonella_ruddii	0	0	0	0	0	0	0	0	0	0	0
Candidatus_nitrospira_defluvii	0	0	0	3.00E-02	2.00E-02	0	0	2.00E-02	2.00E-02	1.00E-02	1.00E-02
Candidatus_portiera_aleyrodidarum	3.00E-02	2.00E-02	1.00E-02	0	0	0.1	0	0	0	0	0
Candidatus_puniceispirillum_marinum	0	0	0	0	0	0	0	0	0	0	0
Candidatus_riesia_pediculicola	0	0	0	0	0	0	0	0	0	0	0
Candidatus_zinderia_insecticola	5.00E-02	9.00E-02	7.00E-02	0	0	0	0	0	0	0	0
Carboxydothermus_hydrogenoformans	0	0	0	0	0	0	0	0	0	0	0
Catenibacterium_mitsuokai	0	0	0	0	0	0	0	0	0	0	0
Cellulophaga_algicola	0	0	0	0	0	0	0	0	0	0	0
Cellulosilyticum_lentocellum	0	0	0	0	0	0	0	0	0	0	0
Chitinophaga_pinensis	0	0	0	0	0	0	0	0	0	0	0
Chlamydia_trachomatis	0	0	0	0	0	0	0	0	0	0	0
Chlamydophila_felis	0	0	0	0	0	0	0	0	0	0	0
Chryseobacterium_gleum	0	0	0	0	0	0	0	0	0	0	0
Clostridiales_genomosp._bvab3	0	0	0	0	1.00E-02	0	2.00E-02	1.00E-02	1.00E-02	2.00E-02	1.00E-02
Clostridium_acidurici	0	0	0	0	0	0	0	0	0	0	0
Clostridium_botulinum	0	0	0	0	0	0	2.00E-02	0	2.00E-02	0	0
Clostridium_cellulovorans	0	0	0	0	0	0	0	0	0	0	0
Clostridium_leptum	0	0	0	0	0	0	0	0	0	0	0
Clostridium_novyi	0	0	0	0	0	0	0	0	0	0	0
Clostridium_perfringens	0	0	0	0	0	0	7.00E-02	0	4.00E-02	0	1.00E-02
Clostridium_sp._sy8519	0	0	0	0	0	0	0	0	0	0	0
Collinsella_aerofaciens	0	0	0	0	0	0	0	0	0	0	0
Comamonas_testosteroni	0	0	0	0	0	0	0	0	0	0	0
Coprococcus_catus	0	0	0	0	0	0	0	0	0	0	0
Coprococcus_comes	0	0	0	0	0	0	0	0	0	0	0
Coprococcus_eutactus	0	0	0	0	0	0	0	0	0	0	0
Coprothermobacter_proteolyticus	1.00E-02	2.00E-02	3.00E-02	0	0	0	0	0	0	0	0
Corynebacterium_accolens	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_ammoniagenes	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_amycolatum	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_aurimucosum	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_glucuronolyticum	0	0	0	0	0	0	2.00E-02	0	0	0	0
Corynebacterium_glutamicum	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_jeikeium	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_kroppenstedtii	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_lipophiloflavum	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_matruchotii	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_pseudogenitalium	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_striatum	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_tuberculostearicum	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_ulcerans	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_urealyticum	0	0	0	0	0	0	0	0	0	0	0
Cronobacter_sakazakii	0	0	0	0	0	0	0	0	0	0	0
Cupriavidus_necator	0	0	0	0	0	0	0	0	0	0	0
Cytophaga_hutchinsonii	0	0	0	0	0	0	0	0	0	0	0
Deinococcus_proteolyticus	0	0	0	0	0	0	0	0	0	0	0
Delftia_acidovorans	0	0	0	0	0	0	0	0	0	0	0
Desulfitobacterium_hafniense	0	0	0	0	0	0	0	0	0	0	0

Table 23. (cont.)

Desulfomonile_tiedjei	0	0	0	0	0	0	0	0	0	0	0
Dialister_invisus	0	0	0	0	0	0	0	0	0	0	0
Dichelobacter_nodosus	0	0	0	0	0	0	0	0	0	0	0
Dickeya_dadantii	0	0	0	0	0	0	0	0	0	0	0
Dorea_longicatena	0	0	0	0	0	0	0	0	0	0	0
Enterobacter_asburiae	0	0	0	0	0	0	0	0	0	0	0
Enterobacter_cloacae	0	0	0	0	0	0	0	0	0	0	0
Enterobacter_sp._638	0	0	0	0	0	0	0	0	0	0	0
Enterococcus_faecalis	0	0	0	0	0	0	0	0	0	0	0
Enterococcus_faecium	0	0	0	0	0	0	0	0	0	0	0
Erysipelothrix_rhusiopathiae	0	0	0	0	0	0	0	0	0	0	0
Erythrobacter_litoralis	0	0	0	0	0	0	2.00E-02	0	0	0	0
Escherichia_coli	2.00E-02	1.00E-02	4.00E-02	0	3.00E-02	0.1	4.00E-02	0	0	0	0
Escherichia_fergusonii	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_eligens	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_hallii	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_limosum	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_rectale	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_siraeum	0	0	0	0	0	0	0	0	0	0	0
Faecalibacterium_prausnitzii	0	0	0	0	0	0	0	0	0	0	0
Ferroglobus_placidus	0	0	0	0	0	0	0	0	0	0	0
Fervidobacterium_pennivorans	0	0	0	0	0	0	0	0	0	0	0
Filifactor_alocis	6.00E-02	7.00E-02	0.11	0	0	0	0	0	0	0	0
Finegoldia_magna	0	0	0	0	1.00E-02	1.25	0	0	1.00E-02	0	1.00E-02
Flavobacterium_johnsoniae	0	0	0	0	0	0	0	0	0	0	0
Fusobacterium_mortiferum	0	0	0	0	0	0	0	0	0	0	0
Fusobacterium_nucleatum	0	0	0	0	0	0	0	0	0	0	1.00E-02
Fusobacterium_periodonticum	0	0	0	0	0	0	0	0	0	0	0
Gardnerella_vaginalis	9.00E-02	1.00E-02	0	0	0	0.29	0	0	0	0	3.00E-02
Gemella_haemolysans	0	0	0	0	0	0	0	0	0	0	0
Gemella_morbillorum	0	0	0	0	0	0	0	0	0	0	0
Gemmatimonas_aurantiaca	0	0	0	0	0	0	0	0	0	0	0
Geobacter_uraniireducens	0	0	0	0	0	0	0	0	0	0	0
Haemophilus_influenzae	0	0	0	1.00E-02	1.00E-02	0	0	0	0	0	0
Haemophilus_parainfluenzae	0	0	0	0	0	0	0	0	0	0	0
Halothermothrix_orenii	0	0	0	0	0	0	0	0	0	0	0
Helicobacter_bilis	0	0	0	0	0	0	0	0	0	0	0
Helicobacter_cetorum	0	0	0	0	1.00E-02	0	0	0	0	0	0
Helicobacter_mustelae	0	0	0	0	0	0	0	0	0	0	0
Helicobacter_pylori	0	0	0	1.00E-02	1.00E-02	0.1	0	1.00E-02	1.00E-02	0	0
Hippea_maritima	0	0	0	0	0	0	0	0	0	0	0
Holdemania_filiformis	0	0	0	0	0	0	0	0	0	0	0
Ignavibacterium_album	0	0	0	0	0	0	0	0	0	0	0
Isoptericola_variabilis	0	0	0	0	0	0	0	0	0	0	0
Jannaschia_sp._ccs1	0	0	0	0	0	0	0	0	0	0	0
Jonquetella_anthropi	0	0	0	0	0	0	0	0	0	0	0
Kangiella_koreensis	0	0	0	0	0	0	0	0	0	0	0
Kineococcus_radiotolerans	0	0	0	0	0	0	0	0	0	0	0
Kitasatospora_setae	0	0	0	0	0	0	0	0	0	0	0
Klebsiella_oxytoca	0	0	0	0	0	0	0	0	0	0	0
Klebsiella_pneumoniae	0	0	0	0	0	0	0	0	0	0	0
Kribbella_flavida	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_acidophilus	0	0	1.00E-02	0.14	0.13	0	4.00E-02	5.00E-02	4.00E-02	5.00E-02	6.00E-02
Lactobacillus_amylolyticus	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_amylovorus	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_antri	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_casei	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_coleohominis	0	0	0	2.00E-02	7.00E-02	0	2.00E-02	1.00E-02	2.00E-02	1.00E-02	3.00E-02
Lactobacillus_crispatus	0	1.00E-02	0.47	86.66	86.62	0.29	80.55	86.78	86.93	86.68	84.71
Lactobacillus_delbrueckii	1.00E-02	0	0	3.00E-02	3.00E-02	0	4.00E-02	3.00E-02	6.00E-02	5.00E-02	3.00E-02
Lactobacillus_fermentum	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_gasseri	2.00E-02	3.00E-02	2.00E-02	1.00E-02	4.00E-02	0.39	0	0	0	0	3.00E-02
Lactobacillus_helveticus	0	1.00E-02	1.00E-02	1.7	1.82	0.1	1.83	1.86	1.89	1.79	1.77
Lactobacillus_iners	89.1	88.72	88.78	0.14	0.39	77.82	4.00E-02	6.00E-02	6.00E-02	5.00E-02	6.00E-02
Lactobacillus_jensenii	0	0	1.00E-02	4.00E-02	4.00E-02	0	7.5	0.1	5.00E-02	0.37	2.28

Table 23. (cont.)

Lactobacillus_johnsonii	0	0	0	0	0	0.1	0	0	0	0	0
Lactobacillus_kefiranofaciens	0	0	0	0	1.00E-02	0	0	0	0	0	0
Lactobacillus_plantarum	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_reuteri	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_rhamnosus	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_sakei	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_ultunensis	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_vaginalis	0	0	0	0	0	0	5.00E-02	7.00E-02	0.11	5.00E-02	0.12
Lactococcus_garvieae	0	0	0	0	0	0	0	0	0	0	0
Lactococcus_lactis	0	0	0	1.00E-02	1.00E-02	0	0	0	0	0	0
Lawsonia_intracellularis	0	0	0	0	0	0	0	0	0	0	0
Leadbetterella_byssophila	0	0	0	0	0	0	0	0	0	0	0
Lysinibacillus_sphaericus	0	0	0	4.00E-02	4.00E-02	0	2.00E-02	6.00E-02	4.00E-02	4.00E-02	4.00E-02
Mahella_australiensis	0	0	0	0	0	0	0	0	0	0	0
Mannheimia_haemolytica	0	0	0	0	0	0	0	0	0	0	0
Marvinbryantia_formatexigens	0	0	0	0	0	0	0	0	0	0	0
Mesoplasma_florum	0	0	0	0	0	0	0	0	0	0	0
Methanobrevibacter_ruminantium	0	0	0	0	0	0	0	0	0	0	0
Methylacidiphilum_infernorum	0	0	0	0	0	0	0	0	0	0	0
Methylomicrobium_alcaliphilum	0	0	0	0	0	0	0	0	0	0	0
Micrococcus_luteus	0	0	0	0	0	0	0	0	0	0	0
Microlunatus_phosphovorus	0	0	0	0	0	0	0	0	0	0	0
Micromonospora_sp._l5	0	0	0	0	0	0	0	0	0	0	0
Mobiluncus_curtisii	0	0	0	0	0	0	0	0	0	0	0
Mycobacterium_abscessus	0	0	0	0	0	0	0	0	0	0	0
Mycobacterium_chubuense	0	0	0	0	0	0	0	0	0	0	0
Mycobacterium_smegmatis	0	0	0	0	0	0	0	0	0	0	0
Mycobacterium_vanbaalenii	0	0	0	0	0	0	0	0	0	0	0
Mycoplasma_arthritidis	0	0	0	0	0	0	0	0	0	0	0
Mycoplasma_crocodyli	0	1.00E-02	1.00E-02	0	0	0	0	0	0	0	0
Mycoplasma_gallisepticum	0	0	0	3.00E-02	1.00E-02	0	0	1.00E-02	2.00E-02	2.00E-02	1.00E-02
Mycoplasma_haemofelis	0	0	0	0	0	0	0	0	0	0	0
Mycoplasma_hominis	0	2.00E-02	1.00E-02	0	0	0	0	0	0	0	0
Mycoplasma_hyopneumoniae	0.17	0.11	0.16	0	0	0.19	0	0	0	0	0
Mycoplasma_synoviae	0	0	0	0	0	0	0	0	0	0	0
Nanoarchaeum_equitans	0	0	0	0	0	0	0	0	0	0	0
Oenococcus_oeni	0	0	0	0	0	0	0	0	0	0	0
Opitutus_terrae	0	0	0	0	0	0	0	0	0	0	0
Orientia_tsutsugamushi	0	0	0	0	0	0	0	0	0	0	0
Oscillibacter_valericigenes	0	0	0	0	0	0	0	0	0	0	0
Owenweeksia_hongkongensis	0	0	0	0	0	0	0	0	0	0	0
Pantoea_ananatis	0	0	0	0	0	0	0	0	0	0	0
Parabacteroides_distasonis	0	0	0	0	0	0	0	0	0	0	0
Parabacteroides_johnsonii	0	0	0	0	0	0	0	0	0	0	0
Parabacteroides_merdae	0	0	0	0	0	0	0	0	0	0	0
Parascardovia_denticolens	0	0	0	0	0	0	0	0	0	0	0
Parvimonas_micra	0	0	0	0	0	0	0	0	0	0	0
Pasteurella_multocida	0	0	0	0	0	0	0	0	0	0	0
Peptoniphilus_lacrimalis	0	0	0	0	0	0	0	0	0	0	0
Peptoniphilus_sp._oral_taxon_386	0	0	0	0	0	0	0	0	0	0	0
Peptostreptococcus_anaerobius	0	0	0	0	0	0	0	0	0	0	0
Peptostreptococcus_stomatis	0	0	0	0	0	0	0	0	0	0	0
Phycisphaera_mikurensis	0	0	0	0	0	0	0	0	0	0	0
Polynucleobacter_necessarius	0	0	0	0	0	0	0	0	0	0	0
Porphyromonas_asaccharolytica	0	0	0	0	0	0	0	0	0	0	0
Porphyromonas_uenonis	0	0	0	0	0	0	0	0	0	0	0
Prevotella_amnii	0	0	0	0	0	0	0	0	0	0	0
Prevotella_bergensis	0	0	0	0	0	0	0	0	0	0	0
Prevotella_bivia	0	0	0	0	0	0	0	0	0	0	0
Prevotella_buccalis	0	0	0	0	0	0	0	0	0	0	0
Prevotella_copri	0	0	0	0	0	0	0	0	0	0	0
Prevotella_melaninogenica	0	0	0	0	0	0	0	0	0	0	0
Prevotella_ruminicola	0	0	0	0	0	0	0	0	0	0	0
Prevotella_sp._oral_taxon_299	0	0	0	0	0	0	0	0	0	0	0
Prevotella_timonensis	0	0	0	0	0	0.96	0	0	0	0	0

Table 23. (cont.)

Prochlorococcus_marinus	0	0	0	1.00E-02	1.00E-02	0	0	1.00E-02	1.00E-02	2.00E-02	1.00E-02
Propionibacterium_acnes	0	0	0	0	0	0	0	0	0	0	1.00E-02
Propionibacterium_freudenreichii	0	0	0	0	0	0	0	0	0	0	0
Pseudoalteromonas_haloplanktis	0	0	0	0	0	0	0	0	0	0	0
Pseudomonas_aeruginosa	0	0	0	0	0	0	0	0	0	0	0
Pseudomonas_fluorescens	0	0	0	0	0	0	0	0	0	0	0
Pseudomonas_mendocina	0	0	0	0	0	0.1	0	0	0	0	0
Pseudomonas_putida	0	0	0	0	0	0	0	0	0	0	0
Pseudomonas_stutzeri	0	0	0	0	0	0	0	0	0	0	0
Psychromonas_ingrahamii	0	0	0	0	0	0	0	0	0	0	0
Pyramidobacter_piscolens	0	0	0	0	0	0	0	0	0	0	0
Rhizobium_leguminosarum	0	0	0	0	0	0	0	0	0	0	0
Rhodococcus_equi	0	0	0	0	0	0	0	0	0	0	0
Rhodococcus_erythropolis	0	0	0	0	0	0	0	0	0	0	0
Rhodopseudomonas_palustris	0	0	0	0	0	0	0	0	0	0	0
Roseburia_hominis	0	0	0	0	0	0	0	0	0	0	0
Roseburia_intestinalis	0	0	0	0	0	0	0	0	0	0	0
Roseburia_inulinivorans	0	0	0	0	0	0	0	0	0	0	0
Rothia_mucilaginosa	0	0	0	0	0	0	0	0	0	0	0
Ruminococcus_albus	0	0	0	0	0	0	0	0	0	0	0
Ruminococcus_champanellensis	0	0	0	0	0	0	0	0	0	0	0
Ruminococcus_lactaris	0	0	0	0	0	0	0	0	0	0	0
Ruminococcus_sp._sr1/5	0	0	0	0	0	0	0	0	0	0	0
Salinibacter_ruber	0	0	0	0	0	0	0	0	0	0	0
Salmonella_enterica	0	0	0	0	0	0	0	0	0	0	0
Scardovia_inopinata	0	0	0	0	0	0	0	0	0	0	0
Selenomonas_ruminantium	0	0	0	0	0	0	0	0	0	0	0
Serratia_symbiotica	0	0	0	0	0	0	0	0	0	0	0
Shewanella_baltica	0	0	0	0	0	0	0	0	0	0	0
Shigella_boydii	0	0	0	0	0	0	0	0	0	0	0
Shigella_dysenteriae	0	0	0	0	0	0	0	0	0	0	0
Shigella_flexneri	0	0	0	0	0	0	0	0	0	0	0
Shigella_sonnei	0	0	0	0	0	0	0	0	0	0	0
Shuttleworthia_satelles	0	0	0	0	0	0	0	0	0	0	0
Simkania_negevensis	0	0	0	0	0	0.1	0	0	0	0	0
Slackia_heliotrinireducens	0	0	0	0	0	0	0	0	0	0	0
Sorangium_cellulosum	0	0	0	0	0	0	0	0	0	0	0
Sphingomonas_wittichii	0	0	0	0	0	0	0	0	0	0	0
Sphingopyxis_alaskensis	0	0	0	0	0	0	0	0	0	0	0
Spirochaeta_africana	0	0	0	0	0	0	0	0	0	0	0
Stackebrandtia_nassauensis	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_aureus	0	0	0	0	0	0	0	0	1.00E-02	0	0
Staphylococcus_epidermidis	0	0	0	1.00E-02	1.00E-02	0	0	2.00E-02	2.00E-02	1.00E-02	0
Staphylococcus_haemolyticus	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_hominis	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_lugdunensis	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_pseudintermedius	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_saprophyticus	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_warneri	0	0	0	0	0	0	0	0	0	0	0
Streptobacillus_moniliformis	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_agalactiae	0	0	0	0	0	0.1	0	0	0	0	0
Streptococcus_anginosus	0	0	0	0	0	1.25	0	0	0	0	1.00E-02
Streptococcus_dysgalactiae	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_equi	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_equinus	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_gallolyticus	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_infantarius	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_intermedius	0	0	0	0	0	0.1	0	0	0	0	0
Streptococcus_mitis	0	0	0	0	0	0	0	1.00E-02	0	0	0
Streptococcus_mutans	0	0	0	0	0	0.1	0	0	0	0	0
Streptococcus_oralis	0	0	0	0	0	0.1	0	0	0	0	0
Streptococcus_parasanguinis	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_parauberis	0	1.00E-02	1.00E-02	0	0	0	0	0	0	0	0
Streptococcus_pneumoniae	0	0	0	0	0	0	0	1.00E-02	0	0	0
Streptococcus_pseudopneumoniae	0	0	0	0	0	0	0	0	0	0	0

Table 23. (cont.)

Streptococcus_pyogenes	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_salivarius	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_sanguinis	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_suis	0	0	1.00E-02	2.00E-02	1.00E-02	0	2.00E-02	3.00E-02	2.00E-02	2.00E-02	1.00E-02
Streptococcus_thermophilus	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_uberis	0	0	0	0	0	0	0	0	0	0	0
Subdoligranulum_variabile	0	0	0	0	0	0	0	0	0	0	0
Sulfolobus_acidocaldarius	0	0	0	0	0	0	0	0	0	0	0
Sulfolobus_islandicus	0	0	0	0	0	0	0	0	0	0	0
Sulfolobus_solfataricus	0	0	0	0	0	0	2.00E-02	0	0	0	0
Sulfolobus_tokodaii	0	0	0	0	0	0	0	0	0	0	0
Sutterella_wadsworthensis	0	0	0	0	0	0	0	0	0	0	0
Synechocystis_sp._pcc_6803	0	0	0	0	0	0	0	0	0	0	0
Teredinibacter_turnerae	0	0	0	0	0	0	0	0	0	0	0
Tetragenococcus_halophilus	0	0	0	0	0	0	0	0	0	0	0
Thermoanaerobacterium_thermosaccharolyticum	0	0	0	0	0	0	0	0	0	0	0
Thermobispora_bispora	0	0	0	0	0	0	0	0	0	0	0
Thermosediminibacter_oceani	0	0	0	0	0	0	0	0	0	0	0
Thermovirga_lienii	0	0	0	0	0	0	0	0	0	0	0
Turicella_otitidis	0	0	0	0	0	0	0	0	0	0	0
Ureaplasma_parvum	0	0	0	0	0	0	0	0	1.00E-02	0	7.00E-02
Ureaplasma_urealyticum	0	0	0	0	0	0	0	0	0	0	0
Variovorax_paradoxus	0	0	0	2.00E-02	2.00E-02	0	0	2.00E-02	1.00E-02	1.00E-02	1.00E-02
Veillonella_parvula	0	0	0	0	0	0	0	0	0	0	0
Weeksella_virosa	0	0	0	0	0	0	0	0	0	0	0
Wolinella_succinogenes	0	0	0	0	0	0	0	0	0	0	0
Xanthomonas_campestris	0	0	0	0	0	0	0	0	0	0	0
Zymomonas_mobilis	0	0	0	2.00E-02	2.00E-02	0	2.00E-02	2.00E-02	3.00E-02	4.00E-02	5.00E-02

Table 23. (cont.)

Bacterial species	117_6	119_1	119_2	119_3	119_4	119_6	120_1	120_2	120_3	120_4	120_6
[Bacteroides]_pectinophilus	0	0	0	0	0	0.11	0	0	0	0	0
[Clostridium]_bartlettii	0	0	0	0	0	0	0	0	0	0	0
[Clostridium]_sticklandii	0	0	0	0	0	0	0	0	0	0	0
[Ruminococcus]_gnavus	0	0	0	0	0	0	0	0	0	0	0
[Ruminococcus]_obeum	0	0	0	0	0	0	0	0	0	0	0
[Ruminococcus]_torques	0	0	0	0	0	0	0	0	0	0	1.00E-02
Acetobacterium_woodii	0	0	0	0	0	0	0	0	0	0	0
Acholeplasma_laidlawii	0	0	0	0	0	0	0	0	0	0	0
Acidimicrobium_ferrooxidans	0	0	0	0	0	0	0	0	0	0	0
Acidothermus_cellulolyticus	2.00E-02	0	0	0	0	0	0	0	0	0	0
Aciduliprofundum_boonei	0	0	0	0	0	0	0	0	0	0	0
Acinetobacter_baumannii	0	0	0	0	0	0	0	0	0	0	0
Acinetobacter_junii	0	0	0	0	0	0	0	0	0	0	0
Acinetobacter_lwoffii	0	0	0	0	0	0	0	0	0	0	0
Actinobacillus_succinogenes	0	0	0	0	0	0	0	0	0	0	0
Actinomyces_coleocanis	0	0	0	0	0	0	0	0	0	0	0
Actinomyces_urogenitalis	0	0	0	0	0	0	0	0	0	0	0
Aggregatibacter_aphrophilus	0	0	0	0	0	0	0	0	0	0	0
Alistipes_finegoldii	0	0	0	0	0	0	0	0	0	0	0
Alistipes_shahii	0	0	0	0	0	0	0	0	0	0	0
Alkaliphilus_metalliredigens	0	0	0	0	0	0	0	0	0	0	0
Allochromatium_vinosum	0	0	0	0	0	0	0	0	0	0	0
Alteromonas_macleodii	0	0	0	0	0	0	0	0	0	0	0
Alteromonas_sp._sn2	0	0	0	0	0	0	0	0	0	0	0
Amycolatopsis_mediterranei	0	0	0	0	0	0	0	0	0	0	0
Amycolicicoccus_subflavus	0	0	0	0	0	0	0	0	0	0	0
Anaerococcus_hydrogenalis	7.00E-02	0	0	0	0	0	0	0	1.00E-02	2.00E-02	5.00E-02
Anaerococcus_lactolyticus	2.00E-02	0	0	0	0	0.22	0	0	0	1.00E-02	6.00E-02
Anaerococcus_prevotii	2.00E-02	0	0	0	0	8.41	0	0	1.00E-02	1.00E-02	3.00E-02
Anaerococcus_tetradius	0	0	0	0	0	2.82	0	0	0	0	1.00E-02
Anaerococcus_vaginalis	2.00E-02	0	0	0	0	0	0	0	5.00E-02	1.00E-02	5.00E-02
Anaerotruncus_colihominis	0	0	0	0	0	0	0	0	0	0	1.00E-02
Arcobacter_nitrofigilis	0	0	0	0	0	0	0	0	0	0	0
Atopobium_parvulum	2.8	0	1.00E-02	0	0	0	0	0	0	0	0
Atopobium_vaginae	0	0	0	0	0	0	0	0	0	0	0
Azospira_oryzae	0	0	0	0	0	0	0.11	0	0	0	0
Bacillus_cereus	0	0	0	0	0	0	0	2.00E-02	1.00E-02	0	0
Bacillus_clausii	0	0	0	0	0	0	0	0	0	0	0
Bacillus_pseudofirmus	0	0	0	0	0	0	0	0	0	0	0
Bacillus_thuringiensis	0	0	0	0	0	0	0	0	0	0	0
Bacteriovorax_marinus	0	0	0	0	0	0	0	2.00E-02	1.00E-02	0	0
Bacteroides_caccae	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_cellulosilyticus	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_dorei	0	0	0	0	0	0	0.11	0	0	0	0
Bacteroides_fragilis	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_ovatus	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_stercoris	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_thetaiotaomicron	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_uniformis	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_vulgatus	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_adolescentis	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_angulatum	5.00E-02	0	0	0	0	6.00E-02	0	0	0	0	0
Bifidobacterium_animalis	0	0	0	0	0	0.17	0	1.00E-02	0	0	1.00E-02
Bifidobacterium_bifidum	1.34	0	0	0	0	6.00E-02	0	5.00E-02	4.00E-02	0	1.00E-02
Bifidobacterium_breve	1.15	0	0	0	0	0.66	0	0	0	0	0
Bifidobacterium_catenulatum	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_dentium	2.00E-02	0	0	0	0	6.00E-02	0	0	0	0	0
Bifidobacterium_gallicum	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_longum	78.26	0	0.51	0	0	35.36	0	1.00E-02	0	0	0
Bifidobacterium_pseudocatenulatum	5.00E-02	0	1.00E-02	0	0	0	0	0	0	0	0
Bilophila_wadsworthia	2.00E-02	0	0	0	0	0	0	0	0	0	0
Blautia_hydrogenotrophica	0	0	0	0	0	6.00E-02	0	0	0	0	0
Brevibacterium_mcbrellneri	0	0	0	0	0	0	0	0	1.00E-02	0	1.00E-02

Table 23. (cont.)

Buchnera_aphidicola	0	0	0	0	0	0	0	0	1.00E-02	0	0
Burkholderia_glumae	0	0	0	0	0	0	0	0	1.00E-02	0	0
Burkholderia_phytofirmans	0	0	0	0	0	0	0	1.00E-02	1.00E-02	1.00E-02	1.00E-02
Burkholderia_pseudomallei	0	0	0	0	0	0	0	0	0	1.00E-02	3.00E-02
Butyrivibrio_crossotus	0	0	0	0	0	6.00E-02	0	0	0	0	0
Butyrivibrio_fibrisolvens	0	0	0	0	0	6.00E-02	0	0	0	0	0
Butyrivibrio_proteoclasticus	0	0	0	0	0	6.00E-02	0	0	0	0	0
Caldicellulosiruptor_bescii	0	0	0	0	0	6.00E-02	0	0	0	0	0
Campylobacter_fetus	0	0	0	7.00E-02	0	6.00E-02	0	0	0	0	0
Campylobacter_gracilis	2.00E-02	0	0	0	0	0	0	0	0	0	0
Campylobacter_hominis	5.00E-02	0	0	0	0	6.00E-02	0	0	0	0	1.00E-02
Campylobacter_lari	0	0	0	0	0	0	0	0	0	0	0
Campylobacter_upsaliensis	0	0	0	0	0	0	0	0	0	0	0
Candidatus_accumulibacter_phosphatis	5.00E-02	0	0	0	0	6.00E-02	0	0	0	0	0
Candidatus_azobacteroides_pseudotrichonymphae	0	0	0	0	0	0	0	0	1.00E-02	0	0
Candidatus_baumannia_cicadellinicola	0	0	1.00E-02	0	5.00E-02	0	0	0	0	1.00E-02	0
Candidatus_carsonella_ruddii	0	0	0	0	0	0	0	0	0	0	0
Candidatus_nitrospira_defluvii	0	3.00E-02	1.00E-02	7.00E-02	0	0	0	0	0	2.00E-02	0
Candidatus_portiera_aleyrodidarum	0	0	0	0	0	0	0	0	0	0	0
Candidatus_puniceispirillum_marinum	0	0	0	0	0	0	0	0	0	0	0
Candidatus_riesia_pediculicola	0	0	1.00E-02	0	0	0	0	1.00E-02	1.00E-02	0	0
Candidatus_zinderia_insecticola	0	0	0	0	0	0	0	0	0	0	0
Carboxydothermus_hydrogenoformans	0	0	0	0	0	0	0	0	0	0	0
Catenibacterium_mitsuokai	0	0	0	0	0	0	0	0	0	0	0
Cellulophaga_algicola	0	0	0	0	0	6.00E-02	0	0	0	0	0
Cellulosilyticum_lentocellum	0	0	0	0	0	0	0	0	0	0	0
Chitinophaga_pinensis	0	0	0	0	0	0	0	1.00E-02	0	0	0
Chlamydia_trachomatis	0	0	0	0	0	0	0	0	0	0	0
Chlamydophila_felis	0	0	0	0	0	0	0	0	0	0	0
Chryseobacterium_gleum	0	0	0	0	0	0	0	0	0	0	0
Clostridiales_genomosp._bvab3	0	0	0	0	0	0	0	0	0	0	0
Clostridium_acidurici	0	0	0	0	0	6.00E-02	0	0	0	0	0
Clostridium_botulinum	0	0	0	0	0	0	0	0	1.00E-02	0	0
Clostridium_cellulovorans	0	0	0	0	0	0	0	0	0	0	0
Clostridium_leptum	0	0	0	0	0	0	0	0	0	0	0
Clostridium_novyi	0	0	0	0	0	0	0	0	0	0	0
Clostridium_perfringens	0	0	0	0	0	0	0	0	0	0	0
Clostridium_sp._sy8519	0	0	0	0	0	6.00E-02	0	0	0	0	0
Collinsella_aerofaciens	0	0	0	0	0	0	0	0	0	0	0
Comamonas_testosteroni	0	0	0	0	0	0	0	0	0	0	0
Coprococcus_catus	0	0	0	0	0	0	0	0	0	0	0
Coprococcus_comes	0	0	0	0	0	0	0	0	0	0	0
Coprococcus_eutactus	0	0	0	0	0	0	0	0	0	0	0
Coprothermobacter_proteolyticus	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_accolens	0	0	0	0	0	0	0.11	0	0	0	0
Corynebacterium_ammoniagenes	0	0	0	0	0	0	0	0	0	0	1.00E-02
Corynebacterium_amycolatum	2.00E-02	0	0	0	0	0	0	0	1.00E-02	1.00E-02	0.11
Corynebacterium_aurimucosum	0	0	0	0	0	0	0.32	0	0	0	3.00E-02
Corynebacterium_glucuronolyticum	0	0	0	0	0	0	0.11	0	0	0	0.11
Corynebacterium_glutamicum	2.00E-02	0	0	0	0	0	0	0	0	0	0
Corynebacterium_jeikeium	0	0	0	0	0.1	0	0.43	0	1.00E-02	1.00E-02	1.00E-02
Corynebacterium_kroppenstedtii	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_lipophiloflavum	0	0	0	0	0	0	0	0	0	0	1.00E-02
Corynebacterium_matruchotii	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_pseudogenitalium	0	0	0	0	0	0	0	0	0	0	2.00E-02
Corynebacterium_striatum	0	0	0	0	0	0	0.11	0	0	0	1.00E-02
Corynebacterium_tuberculostearicum	0	0	0	0	0	0	0	0	0	1.00E-02	1.00E-02
Corynebacterium_ulcerans	0	0	0	0	0	0	0	1.00E-02	0	0	0
Corynebacterium_urealyticum	0	0	0	0	0	0	0	0	0	0	2.00E-02
Cronobacter_sakazakii	0	0	0	0	0	0	0	0	0	0	4.00E-02
Cupriavidus_necator	0	0	0	0	0	0	0	0	0	0	0
Cytophaga_hutchinsonii	0	0	0	0	0	0	0	3.00E-02	1.00E-02	0	0
Deinococcus_proteolyticus	0	0	0	0	0	0	0	0	0	0	0
Delftia_acidovorans	0	0	0	0	0	0	0	0	0	0	0
Desulfitobacterium_hafniense	0	0	0	0	0	0	0	0	0	0	0

Table 23. (cont.)

Desulfomonile_tiedjei	0	0	0	0	0	0	0	0	0	0	1.00E-02
Dialister_invisus	0	0	1.00E-02	0	0	6.00E-02	0	0	1.00E-02	0	0
Dichelobacter_nodosus	0	0	0	0	0	0	0	0	0	0	0
Dickeya_dadantii	0	0	0	0	0	0	0	0	0	0	3.00E-02
Dorea_longicatena	0	0	0	0	0	0	0	0	0	0	1.00E-02
Enterobacter_asburiae	0	0	0	0	0	0	0	0	0	0	1.00E-02
Enterobacter_cloacae	0	0	0	0	0	0	0	0	0	0	0.19
Enterobacter_sp._638	0	0	0	0	0	0	0	0	0	0	4.00E-02
Enterococcus_faecalis	0	0	0	0	0	0	0	0	0	0	0
Enterococcus_faecium	0	0	0	0	0	0	0	5.00E-02	1.00E-02	0	0
Erysipelothrix_rhusiopathiae	0	0	0	0	0	0	0	0	0	0	0
Erythrobacter_litoralis	0	0	0	0	0	0	0.11	0	0	0	0
Escherichia_coli	2.00E-02	6.00E-02	1.00E-02	0.14	0	0.11	0.43	0	4.00E-02	2.00E-02	58.47
Escherichia_fergusonii	0	0	0	0	0	0	0	0	1.00E-02	0	0.36
Eubacterium_eligens	5.00E-02	0	0	0	0	0.17	0	0	0	0	1.00E-02
Eubacterium_hallii	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_limosum	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_rectale	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_siraeum	0	0	0	0	0	0	0	0	0	0	0
Faecalibacterium_prausnitzii	0	0	0	0	0	0	0	0	0	0	0
Ferroglobus_placidus	0	0	0	0	0	0	0	0	0	0	0
Fervidobacterium_pennivorans	0	0	0	0	0	0	0	0	0	0	0
Filifactor_alocis	0	0	0	0	0	0	0	0	0	0	0
Finegoldia_magna	5.00E-02	0	0	0	0	0	0.32	0	4.00E-02	3.00E-02	0.41
Flavobacterium_johnsoniae	0	0	0	0	0	0	0	0	0	0	0
Fusobacterium_mortiferum	0	0	0	0	0	0	0	0	0	0	0
Fusobacterium_nucleatum	0.97	0	0	0	0	0	0	0	0	0	1.00E-02
Fusobacterium_periodonticum	0	0	0	0	0	0	0	0	0	0	0
Gardnerella_vaginalis	2.00E-02	0	0	7.00E-02	0.55	1	61.28	39.81	53.88	3.59	9.92
Gemella_haemolysans	0	0	0	0	0	0	0	0	0	0	0
Gemella_morbillorum	5.00E-02	0	0	0	0	0	0	0	0	0	0
Gemmatimonas_aurantiaca	7.00E-02	0	0	0	0	0	0	0	0	0	0
Geobacter_uraniireducens	0	0	0	0	0	0	0	0	0	0	0
Haemophilus_influenzae	0	0	0	0	0	0	0	0	0	0	0
Haemophilus_parainfluenzae	0	0	0	0	0	0	0	0	0	0	0
Halothermothrix_orenii	0	0	0	0	0	0	0	0	0	0	0
Helicobacter_bilis	0	0	0	0	0	0	0	0	0	0	0
Helicobacter_cetorum	0	0	0	0	0	0	0	0	0	0	0
Helicobacter_mustelae	0	0	0	0	0	0	0	0	0	0	0
Helicobacter_pylori	0	0	0	0	0	0	0	0	0	0	1.00E-02
Hippea_maritima	0	0	0	0	0	0	0	0	0	0	0
Holdemania_filiformis	0	0	0	0	0	0	0	0	0	0	0
Ignavibacterium_album	0	0	0	0	0	0	0	0	0	0	0
Isoptericola_variabilis	0	0	0	0	0	0	0	0	0	0	0
Jannaschia_sp._ccs1	0	0	0	0	0	0	0	0	0	0	0
Jonquetella_anthropi	0	0	0	0	0	0	0	0	0	0	1.00E-02
Kangiella_koreensis	0	0	0	0	0	0	0	0	0	0	0
Kineococcus_radiotolerans	2.00E-02	0	0	0	0	0	0	0	0	0	0
Kitasatospora_setae	0	0	0	0	0	0	0	0	2.00E-02	0	1.00E-02
Klebsiella_oxytoca	0	0	0	0	0	0	0	0	0	0	0.14
Klebsiella_pneumoniae	0	0	0	0	0	0	0	0	0	0	7.00E-02
Kribbella_flavida	0	0	0	0	0	0	0	0	0	0	1.00E-02
Lactobacillus_acidophilus	0	0	0	7.00E-02	0	0	0	0	0	1.00E-02	0
Lactobacillus_amylolyticus	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_amylovorus	0	0	0	0	0	0	0	0	0	1.00E-02	0
Lactobacillus_antri	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_casei	5.00E-02	0	0	0	0	0	0	0	0	0	0
Lactobacillus_coleohominis	0	0	0	0	0	0	0	0	0	1.00E-02	0
Lactobacillus_crispatus	0.19	85.55	86.75	88.62	81.11	0.33	0.11	7.00E-02	4.00E-02	82.12	3.00E-02
Lactobacillus_delbrueckii	2.00E-02	9.00E-02	4.00E-02	7.00E-02	0	0	0	0	1.00E-02	5.00E-02	0
Lactobacillus_fermentum	0	0	0	0	0	0	0	1.00E-02	0	0	5.00E-02
Lactobacillus_gasseri	5.00E-02	9.00E-02	1.00E-02	0	0.25	0.17	23.1	46.98	33.92	0.41	0.19
Lactobacillus_helveticus	0	1.77	1.87	1.86	1.25	0	0	0	1.00E-02	1.51	0
Lactobacillus_iners	2.00E-02	3.00E-02	8.00E-02	0	0.25	0	0.32	0	4.00E-02	7.00E-02	1.00E-02
Lactobacillus_jensenii	0	9.00E-02	2.00E-02	0	5.00E-02	0	0	0	1.00E-02	3.00E-02	2.00E-02

Table 23. (cont.)

Lactobacillus_johnsonii	0	0	0	0	0	0	2.57	3.59	2.62	3.00E-02	2.00E-02
Lactobacillus_kefiranofaciens	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_plantarum	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_reuteri	0	0	0	0	0	0	0	0	0	1.00E-02	0
Lactobacillus_rhamnosus	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_sakei	0	0	0	0	0	0	0	0	1.00E-02	0	0
Lactobacillus_ultunensis	0	0	0	0	0	0	0	0	0	1.00E-02	0
Lactobacillus_vaginalis	0	0	0	0	0	0	0.11	1.05	1.15	1.36	1.00E-02
Lactococcus_garvieae	0	0	0	0	0	0	0	0	0	0	0
Lactococcus_lactis	0	0	0	0	0	0	0	0	0	3.00E-02	0
Lawsonia_intracellularis	0	0	0	0	0	0	0	0	0	0	0
Leadbetterella_byssophila	0	0	0	0	0	0	0	1.00E-02	3.00E-02	0	0
Lysinibacillus_sphaericus	0	3.00E-02	3.00E-02	0	0	0	0	0	0	9.00E-02	0
Mahella_australiensis	0	0	0	0	0	0	0	0	0	0	0
Mannheimia_haemolytica	0	0	0	0	0	0	0	0	0	0	0
Marvinbryantia_formatexigens	0	0	0	0	0	0	0	0	0	0	0
Mesoplasma_florum	0	0	0	0	0	0	0	0	0	0	0
Methanobrevibacter_ruminantium	0	0	0	0	0	0	0	0	1.00E-02	0	0
Methylacidiphilum_infernorum	0	0	0	0	0	0	0	0	0	0	0
Methylomicrobium_alcaliphilum	0	0	0	0	0	0	0	1.00E-02	3.00E-02	1.00E-02	1.00E-02
Micrococcus_luteus	0	3.00E-02	0	0	0	0	0	0	0	0	0
Microlunatus_phosphovorus	0	0	0	0	0	6.00E-02	0	0	0	0	0
Micromonospora_sp._l5	0	0	0	0	0	0	0	0	0	0	0
Mobiluncus_curtisii	2.00E-02	0	0	0	0	0	0	0	0	1.00E-02	3.00E-02
Mycobacterium_abscessus	0	0	0	0	0	6.00E-02	0	0	0	0	0
Mycobacterium_chubuense	0	0	0	0	0	0	0	0	0	0	0
Mycobacterium_smegmatis	0	0	0	0	0	0	0	0	0	0	0
Mycobacterium_vanbaalenii	0	0	0	0	0	0	0	0	0	0	0
Mycoplasma_arthritidis	0	0	0	0	0	0	0	0	0	0	0
Mycoplasma_crocodyli	0	0	0	0	0	0	0	0	0	0	0
Mycoplasma_gallisepticum	0	3.00E-02	3.00E-02	0	5.00E-02	0	0	1.00E-02	1.00E-02	1.00E-02	1.00E-02
Mycoplasma_haemofelis	0	0	0	0	0	6.00E-02	0	0	0	0	0
Mycoplasma_hominis	0	0	0	0	0	0	0	0	0	0	0
Mycoplasma_hyopneumoniae	0	0	0	0	0	0	0	0	1.00E-02	0	0
Mycoplasma_synoviae	0	0	0	0	0	0	0	0	1.00E-02	0	0
Nanoarchaeum_equitans	0	0	0	0	0	0	0	0	0	0	0
Oenococcus_oeni	0	0	0	0	0	0	0	0	0	0	0
Opitutus_terrae	0	0	0	0	0	0	0	0	0	0	0
Orientia_tsutsugamushi	0	0	0	0	0	0	0	1.00E-02	0	0	0
Oscillibacter_valericigenes	0	0	0	0	0	0	0	0	0	0	0
Owenweeksia_hongkongensis	0	0	0	0	0	0	0	3.00E-02	1.00E-02	0	0
Pantoea_ananatis	0	0	0	0	0	0	0	0	0	0	0
Parabacteroides_distasonis	0	0	0	0	0	0	0	0	0	0	0
Parabacteroides_johnsonii	0	0	0	0	0	0	0	0	0	0	0
Parabacteroides_merdae	0	0	0	0	0	0	0	0	0	0	0
Parascardovia_denticolens	0	0	0	0	0	0	0	0	0	0	0
Parvimonas_micra	5.00E-02	0	0	0	0	0	0	0	0	0	0
Pasteurella_multocida	0	0	0	0	0	0	0	0	0	0	0
Peptoniphilus_lacrimalis	0	0	0	0	0	0	0	0	0	0	0
Peptoniphilus_sp._oral_taxon_386	0	0	0	0	0	0	0	0	0	0	0
Peptostreptococcus_anaerobius	0	0	0	0	5.00E-02	3.49	0	0	0	0	0.11
Peptostreptococcus_stomatis	0	0	0	0	0	0	0	0	0	0	0
Phycisphaera_mikurensis	0	0	0	0	0	0	0	0	0	0	0
Polynucleobacter_necessarius	0	0	0	0	0	0	0	1.00E-02	0	0	0
Porphyromonas_asaccharolytica	0	0	0	0	0	6.00E-02	0	0	0	0	0
Porphyromonas_uenonis	0	0	0	0	0	0	0	0	0	0	1.00E-02
Prevotella_amnii	0	0	0	0	0	6.00E-02	0	0	0	0	0
Prevotella_bergensis	0	0	0	0	0	6.00E-02	0	0	1.00E-02	0	0
Prevotella_bivia	0	3.00E-02	2.00E-02	0	0	0.17	0	0	1.00E-02	0	0
Prevotella_buccalis	0	0	0	7.00E-02	0	0	0	0	1.00E-02	1.00E-02	2.00E-02
Prevotella_copri	0	0	0	0	0	0	0	0	0	0	1.00E-02
Prevotella_melaninogenica	0	0	0	0	0	0	0	0	0	0	0
Prevotella_ruminicola	0	0	0	0	0	0	0	0	0	0	0
Prevotella_sp._oral_taxon_299	0	0	0	0	0	0	0	0	0	0	0
Prevotella_timonensis	0.71	0	4.00E-02	0	0	0.5	0	0	0.13	6.00E-02	0.25

Table 23. (cont.)

Prochlorococcus_marinus	0	0	2.00E-02	0	0	0	0	0	0	1.00E-02	0
Propionibacterium_acnes	0	0	0	0	0	0	0.21	0	1.00E-02	0	1.00E-02
Propionibacterium_freudenreichii	0	0	0	0	0	0	0	0	0	0	0
Pseudoalteromonas_haloplanktis	0	0	0	0	0	0	0	0	0	0	0
Pseudomonas_aeruginosa	0	0	0	0	0	0	0	0	0	0	0
Pseudomonas_fluorescens	0	0	0	0	0	0	0	0	0	0	1.00E-02
Pseudomonas_mendocina	0	0	0	0	0	0	0	0	0	0	0
Pseudomonas_putida	0	0	0	0	0	0	0	0	0	0	0
Pseudomonas_stutzeri	0	0	0	0	0	0	0	0	0	0	1.00E-02
Psychromonas_ingrahamii	0	0	0	0	0	0	0	0	0	0	0
Pyramidobacter_piscolens	0	0	0	0	0	0	0	0	0	0	0
Rhizobium_leguminosarum	0	0	0	0	0	0	0	5.00E-02	6.00E-02	0	0
Rhodococcus_equi	0	0	0	0	0	0	0	0	0	0	0
Rhodococcus_erythropolis	0	0	0	0	0	0	0	0	0	0	0
Rhodopseudomonas_palustris	2.00E-02	0	0	0	0	0	0	0	0	0	0
Roseburia_hominis	0	0	0	0	0	0	0	0	0	0	0
Roseburia_intestinalis	0	0	0	0	0	0	0	0	0	0	0
Roseburia_inulinivorans	0	0	0	0	0	0	0	0	0	0	0
Rothia_mucilaginosa	0	0	0	0	0	0	0	0	0	0	0
Ruminococcus_albus	0	0	0	0	0	0	0	0	0	0	0
Ruminococcus_champanellensis	0	0	0	0	0	0	0	0	0	0	1.00E-02
Ruminococcus_lactaris	0	0	0	0	0	0	0	0	0	0	0
Ruminococcus_sp._sr1/5	0	0	0	0	0	0	0	0	0	0	0
Salinibacter_ruber	0	0	0	0	0	0	0	0	0	0	2.00E-02
Salmonella_enterica	0	0	1.00E-02	0	0	0	0	0	0	0	3.5
Scardovia_inopinata	0	0	0	0	0	6.00E-02	0	0	0	0	0
Selenomonas_ruminantium	0	0	0	0	0	0	0	0	0	0	0
Serratia_symbiotica	0	0	0	0	0	0	0	1.00E-02	1.00E-02	0	0
Shewanella_baltica	0	0	0	0	0	0	0	0	0	1.00E-02	0
Shigella_boydii	0	0	0	0	0	0	0	0	0	0	0
Shigella_dysenteriae	0	0	0	0	0	0	0	0	0	0	1.00E-02
Shigella_flexneri	0	0	0	0	0	0	0	0	1.00E-02	0	1.9
Shigella_sonnei	0	0	0	0	0	0	0	0	0	0	1.00E-02
Shuttleworthia_satelles	0	0	0	0	0	0	0	0	0	0	0
Simkania_negevensis	0	0	0	0	0	0	0	0	0	0	0
Slackia_heliotrinireducens	0	0	0	0	0	6.00E-02	0	0	0	0	0
Sorangium_cellulosum	7.00E-02	0	0	0	0	6.00E-02	0	0	0	0	0
Sphingomonas_wittichii	0	0	0	0	0	0	0.11	0	0	0	0
Sphingopyxis_alaskensis	0	0	0	0	0	0	0	0	0	0	0
Spirochaeta_africana	0	0	0	0	0	0	0	0	0	0	0
Stackebrandtia_nassauensis	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_aureus	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_epidermidis	0	3.00E-02	1.00E-02	0	0	0	0.32	0	0	0	0.12
Staphylococcus_haemolyticus	0	0	0	0	0	0	0	0	0	0	9.00E-02
Staphylococcus_hominis	0	0	0	0	0	0	0	0	0	0	2.00E-02
Staphylococcus_lugdunensis	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_pseudintermedius	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_saprophyticus	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_warneri	0	0	0	0	0	0	0	0	0	0	0
Streptobacillus_moniliformis	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_agalactiae	0	1.23	1.00E-02	0	5.66	6.00E-02	0	0	0	0	1.00E-02
Streptococcus_anginosus	0.19	3.00E-02	3.00E-02	0	0.6	27.17	0	0	4.00E-02	1.00E-02	0.8
Streptococcus_dysgalactiae	0	3.00E-02	0	0	0	0	0	0	0	0	0
Streptococcus_equi	2.00E-02	0	0	0	0	0	0	0	0	0	0
Streptococcus_equinus	0	0	0	0	0	6.00E-02	0	0	0	0	0
Streptococcus_gallolyticus	0	0	0	0	0	0.44	0	0	0	0	0
Streptococcus_infantarius	0	0	0	0	0	0	0	0	0	0	1.00E-02
Streptococcus_intermedius	1.46	0	1.00E-02	0	0	0.72	0	0	0	0	5.00E-02
Streptococcus_mitis	0	0	0	0	0	0.44	0.32	0	0	0	1.00E-02
Streptococcus_mutans	5.00E-02	0	0	0	5.00E-02	6.00E-02	0	1.00E-02	0	0	0
Streptococcus_oralis	0	3.00E-02	0	0	0	0.22	0.21	0	0	0	2.00E-02
Streptococcus_parasanguinis	2.00E-02	0	0	0	0	0.17	0.11	4.00E-02	4.00E-02	0	0
Streptococcus_parauberis	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_pneumoniae	0	0	0	0	0	0	0.11	0	0	0	0
Streptococcus_pseudopneumoniae	0	0	0	0	0	0	0	0	0	0	0

Table 23. (cont.)

Streptococcus_pyogenes	2.00E-02	6.00E-02	0	0	5.00E-02	6.00E-02	0	0	0	0	0
Streptococcus_salivarius	0	0	0	0	0	0	0	0	0	0	1.00E-02
Streptococcus_sanguinis	5.00E-02	0	0	0	0	0.22	0	1.00E-02	0	0	1.00E-02
Streptococcus_suis	0	9.00E-02	3.00E-02	0	0	0	0	0	0	1.00E-02	0
Streptococcus_thermophilus	0	0	0	0	0	0	0	0	0	0	1.00E-02
Streptococcus_uberis	0	0	0	0	0	0	0	0	0	0	0
Subdoligranulum_variabile	0	0	0	0	0	0	0	0	0	0	0
Sulfolobus_acidocaldarius	0	0	0	0	0	0	0	0	0	0	0
Sulfolobus_islandicus	0	0	0	0	0	0	0	0	0	0	2.00E-02
Sulfolobus_solfataricus	0	0	0	0	0	0	0	0	0	0	0
Sulfolobus_tokodaii	0	0	0	0	0	0	0	0	0	0	0
Sutterella_wadsworthensis	0	0	0	0	0	0	0	0	0	1.00E-02	1.00E-02
Synechocystis_sp._pcc_6803	0	0	0	0	0	0	0	0	0	0	1.00E-02
Teredinibacter_turnerae	0	0	0	0	0	0	0	0	0	0	0
Tetragenococcus_halophilus	0	0	0	0	0	0	0	0	0	0	0
Thermoanaerobacterium_thermosaccharolyticum	0	0	0	0	0	0	0	0	0	0	0
Thermobispora_bispora	0	0	0	0	0	0	0	0	0	0	0
Thermosediminibacter_oceani	0	0	0	0	0	0	0	0	0	0	0
Thermovirga_lienii	0	0	0	0	0	0	0	0	0	0	0
Turicella_otitidis	0	0	0	0	0	0	0	0	0	0	0
Ureaplasma_parvum	0	0	0	0	0	0	0	0	0	0	0
Ureaplasma_urealyticum	0	0	0	0	0	0	0	0	0	0	0
Variovorax_paradoxus	0	0	3.00E-02	7.00E-02	0	0	0	0	0	1.00E-02	0
Veillonella_parvula	0	0	0	0	0	0	0	0	0	0	1.00E-02
Weeksella_virosa	0	0	0	0	0	0	0	0	0	0	0
Wolinella_succinogenes	0	0	0	0	0	0	0	0	0	0	0
Xanthomonas_campestris	0	0	0	0	0	0	0	0	0	0	0
Zymomonas_mobilis	0	3.00E-02	1.00E-02	0	5.00E-02	0	0	0	0	3.00E-02	0

Table 23. (cont.)

Bacterial species	125_1	125_2	125_3	125_4	125_6	201_1	201_2	201_3	201_5	202_1	202_2
[Bacteroides]_pectinophilus	0	0	0	0	0	0	0	0	0	0	0
[Clostridium]_bartlettii	0	0	0	0	0	0	0	0	0	1.00E-02	0
[Clostridium]_sticklandii	0	0	0	0	0	0	0	0	0	0	0
[Ruminococcus]_gnavus	0	0	0	0	0	0	0	0	0	1.00E-02	0
[Ruminococcus]_obeum	0	0	0	0	0	0	0	0	0	1.00E-02	0
[Ruminococcus]_torques	0	0	0	0	0	0	0	0	0	0	0
Acetobacterium_woodii	0	0	0	0	0	0	0	0	0	0	0
Acholeplasma_laidlawii	0	0	0	1.00E-02	0	0	0	0	0	0	0
Acidimicrobium_ferrooxidans	0	0	0	0	0	0	0	0	0	0	0
Acidothermus_cellulolyticus	0	0	0	0	0	0	0	0	0	0	0
Aciduliprofundum_boonei	0	0	0	0	0	0	0	0	0	0	0
Acinetobacter_baumannii	0	0	0	0	0	0	1.00E-02	0	0	0	0
Acinetobacter_junii	0	0	0	0	0	0	0	0	0	0	0
Acinetobacter_lwoffii	0	0	0	0	0	0	0	0	0	0	0
Actinobacillus_succinogenes	0	1.00E-02	0	0	0	0	0	0	0	0	0
Actinomyces_coleocanis	0	0	0	0	0	0	0	0	0	0	0
Actinomyces_urogenitalis	0	0	0	0	0	0	0	0	0	0	0
Aggregatibacter_aphrophilus	0	0	0	0	0	0	0	0	0	0	0
Alistipes_finegoldii	0	0	0	0	0	0	0	0	0	0	0
Alistipes_shahii	0	0	0	0	0	0	0	0	0	1.00E-02	0
Alkaliphilus_metalliredigens	0	0	0	0	0	0	0	0	0	0	0
Allochromatium_vinosum	0	0	0	0	0	0	0	0	0	0	0
Alteromonas_macleodii	0	0	0	0	0	0	0	0	0	1.00E-02	0
Alteromonas_sp._sn2	0	0	0	0	2.00E-02	0	0	0	0	0	0
Amycolatopsis_mediterranei	0	0	0	0	0	0	0	0	0	0	0
Amycolicicoccus_subflavus	0	0	0	0	0	0	0	0	0	0	0
Anaerococcus_hydrogenalis	0	1.00E-02	0	0	0.1	0	0	0	0	1.00E-02	0
Anaerococcus_lactolyticus	0	2.00E-02	7.00E-02	0	7.00E-02	1.00E-02	0	0	0	2.00E-02	1.00E-02
Anaerococcus_prevotii	0	1.00E-02	1.00E-02	0	0	0	0	0	0	1.00E-02	0
Anaerococcus_tetradius	0	2.00E-02	0.11	0	0	0	0	0	0	0	0
Anaerococcus_vaginalis	0	1.00E-02	0	0	0.17	0	0	0	0	1.00E-02	0
Anaerotruncus_colihominis	0	0	0	0	0	0	0	0	0	1.00E-02	0
Arcobacter_nitrofigilis	0	0	0	1.00E-02	0	0	0	0	0	0	0
Atopobium_parvulum	0	0	0	0	0	0	0	0	0	0	0
Atopobium_vaginae	1.00E-02	0	0	1.85	90.13	4.00E-02	0	2.00E-02	5.00E-02	0	0
Azospira_oryzae	0	0	0	0	0	0	0	0	0	0	0
Bacillus_cereus	1.00E-02	0	0	0	0	0	0	0	0	4.00E-02	1.00E-02
Bacillus_clausii	0	0	0	0	0	0	0	0	0	0	0
Bacillus_pseudofirmus	0	1.00E-02	0	1.00E-02	0	0	0	0	0	0	0
Bacillus_thuringiensis	0	0	0	0	0	0	1.00E-02	0	0	0	0
Bacteriovorax_marinus	1.00E-02	0	0	0	0	0	0	0	0	1.00E-02	0
Bacteroides_caccae	0	0	0	0	0	0	0	0	0	1.00E-02	0
Bacteroides_cellulosilyticus	0	0	0	0	0	0	0	0	0	2.00E-02	0
Bacteroides_dorei	0	0	0	0	0	0	0	0	0	4.00E-02	0
Bacteroides_fragilis	0	0	0	1.00E-02	0	0	0	0	0	3.00E-02	0
Bacteroides_ovatus	0	0	0	0	0	0	0	0	0	5.00E-02	0
Bacteroides_stercoris	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_thetaiotaomicron	0	0	0	0	0	0	0	0	0	2.00E-02	0
Bacteroides_uniformis	0	0	0	0	0	0	0	0	0	8.00E-02	0
Bacteroides_vulgatus	0	0	0	0	0	0	0	0	0	0.12	0
Bifidobacterium_adolescentis	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_angulatum	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_animalis	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_bifidum	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_breve	0	0	0	0	5.00E-02	0	0	0	0	1.00E-02	0
Bifidobacterium_catenulatum	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_dentium	0	0	0	0	0	0	0	0	0	0	2.00E-02
Bifidobacterium_gallicum	0	0	0	0	0	0	0	0	0	0	3.00E-02
Bifidobacterium_longum	0	0	0	0	0	0	0	0	0	0	1.00E-02
Bifidobacterium_pseudocatenulatum	0	0	0	0	0	0	0	0	0	0	3.00E-02
Bilophila_wadsworthia	0	0	0	0	0	0	0	0	0	0	0
Blautia_hydrogenotrophica	0	0	0	1.00E-02	0	0	0	0	0	0	0
Brevibacterium_mcbrellneri	0	0	0	0	0	0	0	0	0	0	0

Table 23. (cont.)

Buchnera_aphidicola	0	0	0	0	0	0	0	0	1.00E-02	1.00E-02	1.00E-02
Burkholderia_glumae	0	0	0	0	0	0	0	0	0	0	0
Burkholderia_phytofirmans	0	0	0	0	0	0	0	0	0	0	0
Burkholderia_pseudomallei	0	0	0	0	0	0	0	0	0	0	0
Butyrivibrio_crossotus	0	0	0	0	0	0	0	0	0	0	0
Butyrivibrio_fibrisolvens	0	0	0	0	0	0	0	0	0	1.00E-02	0
Butyrivibrio_proteoclasticus	0	0	0	0	0	0	0	0	0	0	0
Caldicellulosiruptor_bescii	0	0	0	0	0	0	0	0	0	0	0
Campylobacter_fetus	0	0	0	0	0	0	0	0	0	0	0
Campylobacter_gracilis	0	0	0	0	0	0	0	0	0	0	0
Campylobacter_hominis	0	0	0	0	0	0	0	0	0	0	0
Campylobacter_lari	0	0	0	0	0	0	0	0	0	0	0
Campylobacter_upsaliensis	0	0	0	0	0	0	0	0	0	0	0
Candidatus_accumulibacter_phosphatis	0	0	0	0	0	0	0	0	0	0	0
Candidatus_azobacteroides_pseudotrichonymphae	1.00E-02	0	0	0	0	0	0	0	0	2.00E-02	0
Candidatus_baumannia_cicadellinicola	0	0	0	0	0	0	0	0	0	0	0
Candidatus_carsonella_ruddii	0	0	0	0	0	1.00E-02	0	0	0	0	0
Candidatus_nitrospira_defluvii	0	0	0	0	0	0	0	0	0	0	0
Candidatus_portiera_aleyrodidarum	3.00E-02	0	0	2.00E-02	0	4.00E-02	2.00E-02	3.00E-02	2.00E-02	0	0
Candidatus_puniceispirillum_marinum	0	0	0	0	0	0	0	0	0	0	0
Candidatus_riesia_pediculicola	0	0	0	0	0	0	0	0	0	0	0
Candidatus_zinderia_insecticola	2.00E-02	7.00E-02	1.00E-02	5.00E-02	0	9.00E-02	6.00E-02	8.00E-02	9.00E-02	0	0
Carboxydothermus_hydrogenoformans	0	0	0	0	0	0	0	0	0	0	0
Catenibacterium_mitsuokai	0	0	0	0	1.00E-02	0	0	0	0	0	0
Cellulophaga_algicola	0	0	0	0	0	0	0	0	0	0	0
Cellulosilyticum_lentocellum	0	0	0	0	0	0	0	0	0	0	0
Chitinophaga_pinensis	0	0	0	0	0	0	0	0	0	0	0
Chlamydia_trachomatis	1.00E-02	3.00E-02	4.00E-02	2.00E-02	0	0	0	0	0	0	0
Chlamydophila_felis	0	0	0	0	0	0	0	0	0	0	7.00E-02
Chryseobacterium_gleum	0	0	0	0	0	0	0	0	0	0	0
Clostridiales_genomosp._bvab3	0	0	0	0	0	0	0	0	0	0	0
Clostridium_acidurici	0	0	0	0	0	0	0	0	0	0	0
Clostridium_botulinum	1.00E-02	0	0	0	1.00E-02	0	0	0	0	1.00E-02	1.00E-02
Clostridium_cellulovorans	0	0	0	0	0	0	0	0	0	0	0
Clostridium_leptum	0	0	0	0	0	0	0	0	0	1.00E-02	0
Clostridium_novyi	0	0	0	0	0	0	0	0	0	0	0
Clostridium_perfringens	0	0	0	0	2.00E-02	0	0	0	0	0	0
Clostridium_sp._sy8519	0	0	0	0	0	0	0	0	0	0	0
Collinsella_aerofaciens	0	0	0	0	0	0	0	0	0	0	0
Comamonas_testosteroni	0	0	0	0	0	0	0	0	0	0	0
Coprococcus_catus	0	0	0	0	0	0	0	0	0	0	0
Coprococcus_comes	0	0	0	0	0	0	0	0	0	0	0
Coprococcus_eutactus	0	0	0	0	0	0	0	0	0	0	0
Coprothermobacter_proteolyticus	3.00E-02	1.00E-02	0	0	0	2.00E-02	0	3.00E-02	2.00E-02	0	0
Corynebacterium_accolens	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_ammoniagenes	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_amycolatum	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_aurimucosum	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_glucuronolyticum	0	0	1.00E-02	0	0	0	0	0	0	0	0
Corynebacterium_glutamicum	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_jeikeium	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_kroppenstedtii	0	1.00E-02	0	0	0	0	0	0	0	0	0
Corynebacterium_lipophiloflavum	1.00E-02	0	0	0	0	0	0	0	0	0	0
Corynebacterium_matruchotii	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_pseudogenitalium	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_striatum	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_tuberculostearicum	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_ulcerans	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_urealyticum	0	0	1.00E-02	0	0	0	0	0	0	0	0
Cronobacter_sakazakii	0	0	0	0	0	0	0	0	0	0	0
Cupriavidus_necator	0	0	0	1.00E-02	0	0	0	0	0	0	0
Cytophaga_hutchinsonii	2.00E-02	0	0	0	0	0	0	0	0	2.00E-02	2.00E-02
Deinococcus_proteolyticus	0	0	0	0	0	0	0	0	0	0	0
Delftia_acidovorans	0	0	0	0	0	0	0	0	0	0	0
Desulfitobacterium_hafniense	0	0	0	0	0	0	0	0	0	0	0

Table 23. (cont.)

Desulfomonile_tiedjei	0	0	0	0	0	0	0	0	0	0	0
Dialister_invisus	0	0	0	0	0	0	0	0	0	4.00E-02	0
Dichelobacter_nodosus	0	0	0	0	0	0	0	0	0	0	0
Dickeya_dadantii	0	0	0	0	0	0	0	0	0	0	0
Dorea_longicatena	0	0	0	0	0	0	0	0	0	1.00E-02	0
Enterobacter_asburiae	0	0	0	0	0	0	0	0	0	0	0
Enterobacter_cloacae	0	0	0	0	0	0	0	0	0	0	0
Enterobacter_sp._638	0	0	0	0	0	0	0	0	0	0	0
Enterococcus_faecalis	0	0	0	1.00E-02	0	0	0	0	0	0	0
Enterococcus_faecium	0	0	0	0	0	0	0	0	0	2.00E-02	2.00E-02
Erysipelothrix_rhusiopathiae	0	0	0	0	0	0	0	0	0	0	0
Erythrobacter_litoralis	0	0	0	0	0	0	0	0	0	0	0
Escherichia_coli	6.00E-02	4.00E-02	0.15	0.51	0	4.00E-02	2.00E-02	1.00E-02	3.00E-02	1.00E-02	2.00E-02
Escherichia_fergusonii	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_eligens	0	0	0	0	1.00E-02	0	0	0	0	1.00E-02	0
Eubacterium_hallii	0	0	0	0	0	0	0	0	0	1.00E-02	0
Eubacterium_limosum	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_rectale	0	0	0	0	0	0	0	0	0	0.35	0
Eubacterium_siraeum	0	0	0	0	0	0	0	0	0	0	0
Faecalibacterium_prausnitzii	0	0	0	0	0	0	0	0	0	0.1	0
Ferroglobus_placidus	1.00E-02	1.00E-02	1.00E-02	1.00E-02	0	0	0	0	0	0	0
Fervidobacterium_pennivorans	1.00E-02	0	0	0	0	0	0	0	0	0	0
Filifactor_alocis	7.00E-02	7.00E-02	0	4.00E-02	3.00E-02	6.00E-02	7.00E-02	0.13	4.00E-02	0	0
Finegoldia_magna	1.00E-02	8.00E-02	1.00E-02	0	3.00E-02	0	0	0	0	0	0
Flavobacterium_johnsoniae	0	0	0	0	0	0	0	0	0	0	0
Fusobacterium_mortiferum	0	0	0	0	0	0	0	0	0	0	1.00E-02
Fusobacterium_nucleatum	0	0	0	0	0	0	0	0	0	0	0
Fusobacterium_periodonticum	0	0	0	1.00E-02	0	0	0	0	0	0	0
Gardnerella_vaginalis	4.00E-02	3.00E-02	0.11	9.64	3.00E-02	4.00E-02	6.00E-02	2.00E-02	4.00E-02	0	57.25
Gemella_haemolysans	0	0	0	0	0	0	0	0	0	0	1.00E-02
Gemella_morbillorum	0	0	0	0	0	0	0	0	0	0	0
Gemmatimonas_aurantiaca	0	0	0	0	0	0	0	0	0	0	0
Geobacter_uraniireducens	0	0	0	0	0	0	0	0	0	0	0
Haemophilus_influenzae	0	0	0	0	0	0	0	0	0	0	0
Haemophilus_parainfluenzae	0	0	0	0	0	0	0	0	0	2.00E-02	0
Halothermothrix_orenii	0	0	0	0	0	0	0	0	0	0	0
Helicobacter_bilis	0	0	0	0	0	0	0	0	0	0	1.00E-02
Helicobacter_cetorum	0	0	0	0	0	0	0	0	0	0	0
Helicobacter_mustelae	0	0	0	0	1.00E-02	0	0	0	0	0	0
Helicobacter_pylori	0	0	0	0	2.00E-02	0	0	0	0	0	0
Hippea_maritima	0	0	0	0	0	0	0	0	0	0	3.00E-02
Holdemania_filiformis	0	0	0	0	0	0	0	0	0	0	0
Ignavibacterium_album	0	0	0	0	0	0	0	0	0	0	0
Isoptericola_variabilis	0	0	0	0	0	0	0	0	0	0	0
Jannaschia_sp._ccs1	0	0	0	0	0	0	0	0	0	0	0
Jonquetella_anthropi	0	0	0	0	4.00E-02	0	0	0	0	0	0
Kangiella_koreensis	1.00E-02	0	1.00E-02	0	0	0	0	0	0	0	0
Kineococcus_radiotolerans	0	0	0	0	0	0	0	0	0	0	0
Kitasatospora_setae	0	0	0	0	0	0	0	0	0	0	0
Klebsiella_oxytoca	0	0	0	0	0	0	0	0	0	0	0
Klebsiella_pneumoniae	0	0	0	0	0	0	0	0	0	0	0
Kribbella_flavida	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_acidophilus	0	0	0	0	0	0	0	0	0	0	1.00E-02
Lactobacillus_amylolyticus	0	0	0	0	0	0	1.00E-02	0	0	0	0
Lactobacillus_amylovorus	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_antri	0	0	0	0	0	0	2.00E-02	0	1.00E-02	0	0
Lactobacillus_casei	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_coleohominis	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_crispatus	5.00E-02	2.00E-02	7.00E-02	6.00E-02	0	0	2.00E-02	0	0	1.00E-02	2.00E-02
Lactobacillus_delbrueckii	0	2.00E-02	0	0	0	1.00E-02	2.00E-02	1.00E-02	1.00E-02	1.00E-02	1.00E-02
Lactobacillus_fermentum	0	1.00E-02	0	0	0	0	0	0	1.00E-02	0	0
Lactobacillus_gasseri	10.86	7.00E-02	0.11	0.35	5.00E-02	8.00E-02	6.00E-02	7.00E-02	0.1	82.08	31.58
Lactobacillus_helveticus	1.00E-02	1.00E-02	6.00E-02	4.00E-02	2.00E-02	1.00E-02	2.00E-02	2.00E-02	1.00E-02	1.00E-02	0
Lactobacillus_iners	55.92	60.95	54.48	52.38	5.00E-02	88.91	88.95	89.04	88.78	1.00E-02	6.00E-02
Lactobacillus_jensenii	22.08	28.22	33.68	23.75	0	0	0	0	0	0	0

Table 23. (cont.)

Lactobacillus_johnsonii	0.79	0	0	7.00E-02	0	0	1.00E-02	1.00E-02	2.00E-02	6.39	2.32
Lactobacillus_kefiranofaciens	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_plantarum	2.00E-02	4.00E-02	4.00E-02	4.00E-02	0	0	0	0	0	2.00E-02	1.00E-02
Lactobacillus_reuteri	0	0	1.00E-02	0	0	0	0	0	1.00E-02	0	0
Lactobacillus_rhamnosus	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_sakei	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_ultunensis	0	0	0	0	0	0	0	0	0	0	1.00E-02
Lactobacillus_vaginalis	0.14	0.55	0.87	1.1	2.00E-02	0	0	0	0	0	0
Lactococcus_garvieae	0	0	0	0	0	0	0	0	0	0	0
Lactococcus_lactis	0	0	0	0	0	0	0	0	0	0	0
Lawsonia_intracellularis	0	0	0	0	0	0	0	0	0	0	1.00E-02
Leadbetterella_byssophila	0	0	0	0	0	0	0	0	0	0	0
Lysinibacillus_sphaericus	0	0	0	0	0	0	0	0	0	0	0
Mahella_australiensis	0	0	0	0	0	0	0	0	0	0	0
Mannheimia_haemolytica	0	0	0	0	0	0	0	0	0	0	0
Marvinbryantia_formatexigens	0	0	0	0	0	0	0	0	0	0	0
Mesoplasma_florum	1.00E-02	1.00E-02	0	0	3.00E-02	0	0	0	0	0	0
Methanobrevibacter_ruminantium	0	2.00E-02	1.00E-02	1.00E-02	0	0	0	0	0	0	1.00E-02
Methylacidiphilum_infernorum	0	0	0	0	0	0	0	0	0	0	3.00E-02
Methylomicrobium_alcaliphilum	0	0	0	0	0	0	0	0	0	0	0
Micrococcus_luteus	0	0	0	0	0	1.00E-02	0	0	0	0	0
Microlunatus_phosphovorus	0	0	0	0	0	0	0	0	0	0	0
Micromonospora_sp._l5	0	0	0	0	0	0	0	0	0	0	0
Mobiluncus_curtisii	0	0	6.00E-02	0	1.00E-02	0	0	0	0	0	1.00E-02
Mycobacterium_abscessus	0	0	0	0	0	0	0	0	0	0	0
Mycobacterium_chubuense	0	0	0	0	0	0	0	0	0	0	0
Mycobacterium_smegmatis	0	0	0	0	0	0	0	0	0	0	0
Mycobacterium_vanbaalenii	0	0	0	0	0	0	0	0	0	0	0
Mycoplasma_arthritidis	1.00E-02	1.00E-02	1.00E-02	1.00E-02	0	0	0	0	0	0	0
Mycoplasma_crocodyli	0	1.00E-02	0	0	0	2.00E-02	3.00E-02	1.00E-02	1.00E-02	0	0
Mycoplasma_gallisepticum	0	0	0	0	0	0	0	0	0	2.00E-02	3.00E-02
Mycoplasma_haemofelis	0	0	0	0	3.00E-02	0	0	0	0	0	0
Mycoplasma_hominis	1.00E-02	0	1.00E-02	1.00E-02	0	2.00E-02	1.00E-02	0	1.00E-02	0	0
Mycoplasma_hyopneumoniae	0.12	8.00E-02	7.00E-02	7.00E-02	0	0.13	0.13	0.15	0.13	1.00E-02	0
Mycoplasma_synoviae	0	0	0	1.00E-02	0	4.00E-02	1.00E-02	4.00E-02	3.00E-02	1.00E-02	2.00E-02
Nanoarchaeum_equitans	0	0	0	0	0	0	0	0	0	0	0
Oenococcus_oeni	0	0	0	0	0	0	0	0	0	0	1.00E-02
Opitutus_terrae	0	0	0	0	0	0	0	0	0	0	0
Orientia_tsutsugamushi	0	0	0	0	0	0	0	0	0	1.00E-02	1.00E-02
Oscillibacter_valericigenes	0	0	0	0	0	0	0	0	0	1.00E-02	0
Owenweeksia_hongkongensis	0	0	0	0	0	0	0	0	0	2.00E-02	1.00E-02
Pantoea_ananatis	0	3.00E-02	1.00E-02	0	0	1.00E-02	0	0	0	0	0
Parabacteroides_distasonis	0	0	0	0	0	0	0	0	0	1.00E-02	0
Parabacteroides_johnsonii	0	0	0	0	0	0	0	0	0	1.00E-02	0
Parabacteroides_merdae	0	0	0	0	0	0	0	0	0	4.00E-02	0
Parascardovia_denticolens	0	0	0	0	0	0	0	0	0	0	6.00E-02
Parvimonas_micra	0	0	0	0	0	0	0	0	0	0	0
Pasteurella_multocida	0	0	0	0	0	0	0	0	0	1.00E-02	0
Peptoniphilus_lacrimalis	0	0	1.00E-02	0	0	0	0	0	0	0	0
Peptoniphilus_sp._oral_taxon_386	0	0	0	0	0	0	0	0	0	0	0
Peptostreptococcus_anaerobius	0	0	0	0	0	0	0	0	0	0	0
Peptostreptococcus_stomatis	0	0	0	0	1.00E-02	0	0	0	0	0	0
Phycisphaera_mikurensis	0	0	0	0	0	0	0	0	0	0	0
Polynucleobacter_necessarius	0	0	0	0	0	0	0	0	0	0	1.00E-02
Porphyromonas_asaccharolytica	0	0	3.00E-02	0	0	0	0	0	0	1.00E-02	0
Porphyromonas_uenonis	0	0	0	0	2.00E-02	0	0	0	0	0	0
Prevotella_amnii	0	0	0	0	0	0	0	0	0	0	0
Prevotella_bergensis	0	0	0	0	2.00E-02	0	0	0	0	0	0
Prevotella_bivia	0	0	0	0	0.11	0	0	0	0	0	0
Prevotella_buccalis	0	0	3.00E-02	0	2.00E-02	0	0	0	0	0	0
Prevotella_copri	0	0	0	0	0	0	0	0	0	0	0
Prevotella_melaninogenica	0	0	0	0	0	0	0	0	0	0	0
Prevotella_ruminicola	0	0	0	0	0	0	0	0	0	0	0
Prevotella_sp._oral_taxon_299	0	0	0	0	0	0	0	0	0	1.00E-02	0
Prevotella_timonensis	1.00E-02	9.00E-02	0.22	0	0.91	0	0	0	0	1.00E-02	1.00E-02

Table 23. (cont.)

Prochlorococcus_marinus	0	0	0	0	0	0	0	0	0	0	0
Propionibacterium_acnes	0	0	0	0	0	0	1.00E-02	1.00E-02	1.00E-02	0	1.00E-02
Propionibacterium_freudenreichii	0	0	0	0	0	0	0	0	0	0	0
Pseudoalteromonas_haloplanktis	0	1.00E-02	0	0	0	0	0	0	0	0	0
Pseudomonas_aeruginosa	0	0	0	0	0	0	0	0	0	0	0
Pseudomonas_fluorescens	0	0	0	0	0	0	0	0	0	1.00E-02	0
Pseudomonas_mendocina	0	0	0	0	0	1.00E-02	1.00E-02	0	0	0	0
Pseudomonas_putida	0	0	0	0	0	0	0	0	0	0	0
Pseudomonas_stutzeri	0	0	0	0	0	0	0	0	0	0	0
Psychromonas_ingrahamii	0	0	0	0	0	0	0	0	0	0	0
Pyramidobacter_piscolens	0	0	0	0	1.00E-02	0	0	0	0	0	0
Rhizobium_leguminosarum	0	0	0	0	2.00E-02	0	0	0	0	0	0
Rhodococcus_equi	0	1.00E-02	0	0	0	0	0	0	0	0	0
Rhodococcus_erythropolis	0	0	0	0	0	0	0	0	0	0	0
Rhodopseudomonas_palustris	0	0	0	0	0	0	0	0	0	0	0
Roseburia_hominis	0	0	0	0	0	0	0	0	0	1.00E-02	0
Roseburia_intestinalis	0	0	0	0	0	0	0	0	0	5.00E-02	0
Roseburia_inulinivorans	0	0	0	0	0	0	0	0	0	3.00E-02	0
Rothia_mucilaginosa	0	0	0	0	0	0	0	0	0	0	0
Ruminococcus_albus	0	0	0	0	0	0	0	0	0	0	0
Ruminococcus_champanellensis	0	0	0	0	0	0	0	0	0	0	0
Ruminococcus_lactaris	0	0	0	0	0	0	0	0	0	1.00E-02	0
Ruminococcus_sp._sr1/5	0	0	0	0	0	0	0	0	0	1.00E-02	0
Salinibacter_ruber	0	0	0	0	0	0	0	0	0	0	0
Salmonella_enterica	2.00E-02	1.00E-02	0	4.00E-02	0	0	0	0	0	0	0
Scardovia_inopinata	0	0	0	0	0	0	0	0	0	0	0
Selenomonas_ruminantium	0	0	0	0	0	0	0	0	0	0	0
Serratia_symbiotica	0	0	0	0	0	0	0	0	0	0	0
Shewanella_baltica	0	0	0	0	0	0	0	0	0	0	0
Shigella_boydii	0	0	0	0	0	0	0	0	0	0	0
Shigella_dysenteriae	0	0	0	0	0	0	0	0	0	0	0
Shigella_flexneri	0	0	1.00E-02	4.00E-02	0	0	0	0	0	0	0
Shigella_sonnei	0	0	0	0	0	0	0	0	0	0	0
Shuttleworthia_satelles	0	0	0	0	0	0	0	0	0	0	0
Simkania_negevensis	0	0	0	0	0	0	0	0	0	0	0
Slackia_heliotrinireducens	0	0	0	0	0	0	0	0	0	0	0
Sorangium_cellulosum	0	0	0	0	0	0	0	0	0	0	0
Sphingomonas_wittichii	0	0	0	0	0	0	0	0	0	0	0
Sphingopyxis_alaskensis	0	0	0	0	0	0	0	0	0	0	0
Spirochaeta_africana	0	0	0	0	0	0	0	0	0	0	0
Stackebrandtia_nassauensis	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_aureus	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_epidermidis	0	0	1.00E-02	0	0	0	0	0	1.00E-02	0	0
Staphylococcus_haemolyticus	0	0	1.00E-02	0	0	0	0	0	0	0	0
Staphylococcus_hominis	0	0	3.00E-02	0	0	0	0	0	0	0	0
Staphylococcus_lugdunensis	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_pseudintermedius	2.00E-02	2.00E-02	1.00E-02	1.00E-02	0	1.00E-02	4.00E-02	4.00E-02	3.00E-02	0	0
Staphylococcus_saprophyticus	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_warneri	0	0	0	0	0	0	0	0	0	0	0
Streptobacillus_moniliformis	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_agalactiae	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_anginosus	1.00E-02	0	0	0	0	0	0	0	0	0.15	6.00E-02
Streptococcus_dysgalactiae	0	0	0	0	0	0	0	0	0	1.00E-02	0
Streptococcus_equi	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_equinus	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_gallolyticus	0	0	0	0	0	0	0	0	0	0	1.00E-02
Streptococcus_infantarius	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_intermedius	1.00E-02	2.00E-02	1.00E-02	0	0	0	0	0	0	0	0
Streptococcus_mitis	0	0	0	0	0	0	0	0	0	1.00E-02	0
Streptococcus_mutans	1.00E-02	0	0	0	0	0	0	0	0	0	0
Streptococcus_oralis	0	0	0	0	0	0	0	0	0	0	3.00E-02
Streptococcus_parasanguinis	0	0	0	0	0	0	0	0	0	6.00E-02	2.00E-02
Streptococcus_parauberis	0	0	1.00E-02	0	0	0	0	0	0	0	0
Streptococcus_pneumoniae	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_pseudopneumoniae	0	0	0	0	0	0	0	0	0	0	0

Table 23. (cont.)

Streptococcus_pyogenes	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_salivarius	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_sanguinis	0	0	0	0	0	0	0	1.00E-02	1.00E-02	1.00E-02	1.00E-02
Streptococcus_suis	2.00E-02	1.00E-02	4.00E-02	4.00E-02	0	0	0	0	0	0	0
Streptococcus_thermophilus	0	0	0	0	0	0	0	0	0	2.00E-02	0
Streptococcus_uberis	0	0	0	0	0	0	0	0	0	0	0
Subdoligranulum_variabile	0	0	0	0	0	0	0	0	0	0	0
Sulfolobus_acidocaldarius	0	0	0	0	0	1.00E-02	2.00E-02	0	1.00E-02	0	0
Sulfolobus_islandicus	0	0	0	0	0	0	0	0	0	0	0
Sulfolobus_solfataricus	0	0	0	0	0	0	0	0	0	0	0
Sulfolobus_tokodaii	0	0	0	0	0	0	0	0	0	0	0
Sutterella_wadsworthensis	0	0	0	0	0	0	0	0	0	0	0
Synechocystis_sp._pcc_6803	0	0	0	0	0	0	0	0	0	0	0
Teredinibacter_turnerae	0	0	0	0	0	0	0	0	0	0	0
Tetragenococcus_halophilus	0	0	0	0	3.00E-02	0	0	1.00E-02	0	0	0
Thermoanaerobacterium_thermosaccharolyticum	0	0	0	0	0	0	0	0	0	1.00E-02	1.00E-02
Thermobispora_bispora	0	0	0	0	0	0	0	0	0	0	0
Thermosediminibacter_oceani	0	0	0	0	0	0	0	0	0	0	0
Thermovirga_lienii	0	0	0	0	0	0	0	0	0	0	2.00E-02
Turicella_otitidis	0	0	0	0	0	0	0	0	0	0	0
Ureaplasma_parvum	0	0	4.00E-02	0.12	0	0	0	0	1.00E-02	0.72	0.32
Ureaplasma_urealyticum	0	0	1.00E-02	0	0	0	0	0	0	1.00E-02	3.00E-02
Variovorax_paradoxus	0	0	0	0	0	0	0	0	0	0	0
Veillonella_parvula	0	0	0	0	0	0	0	0	0	0	0
Weeksella_virosa	0	1.00E-02	0	1.00E-02	0	0	0	0	0	0	0
Wolinella_succinogenes	0	0	0	0	0	1.00E-02	0	1.00E-02	2.00E-02	0	0
Xanthomonas_campestris	0	0	0	0	0	0	1.00E-02	0	0	1.00E-02	0
Zymomonas_mobilis	0	0	0	0	0	0	0	0	0	0	0

Table 23. (cont.)

Bacterial species	202_3	202_4	202_6	203_1	203_2	203_3	203_4	203_5	203_6	205_1	205_2
[Bacteroides]_pectinophilus	0	0	0	0	0	0	0	2.00E-02	0	0	0
[Clostridium]_bartlettii	0	0	0	0	0	0	0	0	0	0	0
[Clostridium]_sticklandii	0	0	0.15	0	0	0	0	1.00E-02	0	0	0
[Ruminococcus]_gnavus	0	0	0	0	0	0	0	0	0	0	0
[Ruminococcus]_obeum	0	0	0	0	0	0	0	0	0	0	0
[Ruminococcus]_torques	0	0	0	0	0	0	0	0	0	0	0
Acetobacterium_woodii	0	0	0	0	0	0	0	0	0.14	0	0
Acholeplasma_laidlawii	0	0	0	0	3.00E-02	0	0	1.00E-02	0	0	0
Acidimicrobium_ferrooxidans	0	0	0	2.00E-02	0	0	0	0	0	0	0
Acidothermus_cellulolyticus	0	0	0	0	0	0	0	0	0	0	0
Aciduliprofundum_boonei	0	0	0	0	0	0	0	1.00E-02	0	0	0
Acinetobacter_baumannii	0	0	0	0	0	0	0	0	0	0	0
Acinetobacter_junii	0	0	0	0	0	0	0	0	0	0	0
Acinetobacter_lwoffii	1.00E-02	0	0	0	0	0	0	0	0	0	0
Actinobacillus_succinogenes	0	0	0	0	0	0	0	0	0	0	0
Actinomyces_coleocanis	0	0	0.15	0	0	0	0	0	0.14	0	0
Actinomyces_urogenitalis	0	0	0	0	0	0	0	0	0	0	0
Aggregatibacter_aphrophilus	0	0	0	0	0	0	0	4.00E-02	0	0	0
Alistipes_finegoldii	0	0	0.15	0	0	0	0	0	0	0	0
Alistipes_shahii	0	0	0	0	0	0	0	0	0	0	0
Alkaliphilus_metalliredigens	0	0	0.15	0	0	0	0	0	0	0	0
Allochromatium_vinosum	0	0	0	0	0	0	0	1.00E-02	0	0	0
Alteromonas_macleodii	1.00E-02	1.00E-02	0	0	0	0	0	1.00E-02	0	0	0
Alteromonas_sp._sn2	0	0	0	0	0	0	0	0	0	0	0
Amycolatopsis_mediterranei	0	0	0	0	0	0	0	1.00E-02	0	0	0
Amycolicicoccus_subflavus	0	0	0.15	0	0	0	0	0	0	0	0
Anaerococcus_hydrogenalis	0	0	1.2	0	3.00E-02	0	0	0.58	0.69	0	0
Anaerococcus_lactolyticus	1.00E-02	1.00E-02	9.27	0	0	0	0	0.41	29.08	0	0
Anaerococcus_prevotii	0	0	0.15	0	0	0	0	2.00E-02	0.27	0	1.00E-02
Anaerococcus_tetradius	0	0	0.15	0	0	0	0	0	0	0	0
Anaerococcus_vaginalis	0	0	2.69	0	0	0	0	2.00E-02	2.06	0	0
Anaerotruncus_colihominis	0	0	0	0	0	0	0	0	0	0	0
Arcobacter_nitrofigilis	0	0	0	0	0	0	0	0	0	0	0
Atopobium_parvulum	0	0	0	0	0	0	0	0	0	0	0
Atopobium_vaginae	3.00E-02	5.00E-02	0	0	0	0	0	0	0.27	0	0
Azospira_oryzae	0	0	0	0	0	0	0	0	0	0	0
Bacillus_cereus	1.00E-02	1.00E-02	0	0	0	0	0	0	0.14	0	0
Bacillus_clausii	0	0	0	0	0	0	0	0	0.14	0	0
Bacillus_pseudofirmus	0	0	0	0	0	0	0	0	0	0	0
Bacillus_thuringiensis	0	0	0	0	0	0	0	0	0	0	0
Bacteriovorax_marinus	0	2.00E-02	0	0	0	0	0	0	0.14	0	0
Bacteroides_caccae	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_cellulosilyticus	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_dorei	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_fragilis	0	0	0.15	0	0	0	0	2.00E-02	0	0	0
Bacteroides_ovatus	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_stercoris	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_thetaiotaomicron	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_uniformis	0	0	0	0	0	0	0	0	0	0	0
Bacteroides_vulgatus	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_adolescentis	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_angulatum	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_animalis	0	0	0	0	2.00E-02	0	0	0	0.14	0	0
Bifidobacterium_bifidum	0	0	0	9.00E-02	5.00E-02	0	0	6.00E-02	0	0	0
Bifidobacterium_breve	0	0	0	0	0	0	0	2.00E-02	0	0	0
Bifidobacterium_catenulatum	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_dentium	4.00E-02	2.00E-02	0	0	0	0	0	0	0	0	0
Bifidobacterium_gallicum	1.00E-02	1.00E-02	0	0	0	0	0	0	0	0	0
Bifidobacterium_longum	1.00E-02	1.00E-02	0	9.00E-02	8.00E-02	0	0.1	2.04	0	0	0
Bifidobacterium_pseudocatenulatum	3.00E-02	2.00E-02	0	0	0	0	0	0	0	0	0
Bilophila_wadsworthia	0	0	0	0	0	0	0	0	0	0	0
Blautia_hydrogenotrophica	0	0	0	0	0	0	0	1.00E-02	0	0	0
Brevibacterium_mcbrellneri	0	0	0	0	0	0	0	0	0	0	0

Table 23. (cont.)

Buchnera_aphidicola	0	0	0	0	0	0	0	0	0	0	0
Burkholderia_glumae	0	0	0	0	0	0	0	0	0	0	0
Burkholderia_phytofirmans	0	0	0	0	5.00E-02	0	0	0	0	0	0
Burkholderia_pseudomallei	0	0	0	0	0	0	0	0	0	0	0
Butyrivibrio_crossotus	0	0	0	0	0	0	0	0	0	0	0
Butyrivibrio_fibrisolvens	0	0	0	0	0	0	0	0	0	0	0
Butyrivibrio_proteoclasticus	0	0	0	0	0	0	0	0	0	0	0
Caldicellulosiruptor_bescii	0	0	0	0	0	0	0	0	0	0	0
Campylobacter_fetus	0	0	0.75	0	0	0	0	0	0	0	0
Campylobacter_gracilis	0	0	0	0	0	0	0	0	0	0	0
Campylobacter_hominis	0	0	0.9	0	0	0	0	0	0.14	0	0
Campylobacter_lari	0	0	0.15	0	0	0	0	0	0	0	0
Campylobacter_upsaliensis	0	0	0	0	0	0	0	0	0	0	0
Candidatus_accumulibacter_phosphatis	0	0	0	0	0	0	0	0	0	0	0
Candidatus_azobacteroides_pseudotrichonymphae	0	1.00E-02	0	0	0	0	0	0	0	0	0
Candidatus_baumannia_cicadellinicola	0	0	0	0	0	0	0	0	0	2.00E-02	1.00E-02
Candidatus_carsonella_ruddii	0	0	0	0	0	0	0	0	0	0	0
Candidatus_nitrospira_defluvii	0	0	0	0	0	0	0	0	0	4.00E-02	2.00E-02
Candidatus_portiera_aleyrodidarum	0	0	0	0	0	0	0	0	0	0	0
Candidatus_puniceispirillum_marinum	0	0	0	0	0	0	0	0	0	0	0
Candidatus_riesia_pediculicola	1.00E-02	0	0	0	0	0	4.00E-02	0	0	0	0
Candidatus_zinderia_insecticola	0	0	0	0	0	0	0	0	0	0	0
Carboxydothermus_hydrogenoformans	0	0	0	0	0	0	0	0	0	0	0
Catenibacterium_mitsuokai	0	0	0	0	0	0	0	0	0	0	0
Cellulophaga_algicola	0	0	0	0	0	0	0	0	0	0	0
Cellulosilyticum_lentocellum	0	0	0	0	0	0	0	1.00E-02	0.14	0	0
Chitinophaga_pinensis	0	0	0	0	0	0	0	0	0	0	0
Chlamydia_trachomatis	0	0	0	0	0	0	0	0	0	0	0
Chlamydophila_felis	8.00E-02	7.00E-02	0	0	0	0	0	0	0	0	0
Chryseobacterium_gleum	0	0	0	0	0	0	0	0	0	0	0
Clostridiales_genomosp._bvab3	0	0	0	0	0	0	0	0	0	2.00E-02	1.00E-02
Clostridium_acidurici	0	0	0.15	0	0	0	0	0	0.69	0	0
Clostridium_botulinum	1.00E-02	2.00E-02	0.3	0	0	0	0	0	0.27	0	0
Clostridium_cellulovorans	0	0	0	0	0	0	0	0	0.14	0	0
Clostridium_leptum	0	0	0	0	0	0	0	0	0	0	0
Clostridium_novyi	0	0	0	0	0	0	0	0	0.14	0	0
Clostridium_perfringens	0	0	0	0	0	0	0	0	0.27	0	0
Clostridium_sp._sy8519	0	0	0	0	0	0	0	0	0	0	0
Collinsella_aerofaciens	0	0	0	0	0	0	0	0	0	0	0
Comamonas_testosteroni	0	0	0	0	0	0	0	0	0	0	0
Coprococcus_catus	0	0	0	0	0	0	0	1.00E-02	0	0	0
Coprococcus_comes	0	0	0	0	0	0	0	0	0	0	0
Coprococcus_eutactus	0	0	0.15	0	0	0	0	1.00E-02	0	0	0
Coprothermobacter_proteolyticus	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_accolens	0	0	0	0	0	0	0	1.00E-02	0	0	0
Corynebacterium_ammoniagenes	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_amycolatum	0	0	0.45	0	0	0	0	0	0	0	0
Corynebacterium_aurimucosum	0	0	0	0	0	0	0	2.00E-02	0.14	0	1.00E-02
Corynebacterium_glucuronolyticum	0	0	0.15	0	0	0	0	6.00E-02	0	0	0
Corynebacterium_glutamicum	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_jeikeium	0	0	0	0	0	0	0	0.17	0	0	2.00E-02
Corynebacterium_kroppenstedtii	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_lipophiloflavum	0	0	0.3	0	0	0	0	5.00E-02	0	0	0
Corynebacterium_matruchotii	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_pseudogenitalium	0	0	0	0	0	0	0	5.00E-02	0.14	0	1.00E-02
Corynebacterium_striatum	0	0	0	0	0	0	0	1.00E-02	0	0	0
Corynebacterium_tuberculostearicum	0	0	0	0	0	0	0	4.00E-02	0	0	0
Corynebacterium_ulcerans	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_urealyticum	0	0	0	0	0	0	0	6.00E-02	0	0	0
Cronobacter_sakazakii	0	0	0	0	0	0	0	0	0	0	0
Cupriavidus_necator	0	0	0	0	0	0	0	0	0	0	0
Cytophaga_hutchinsonii	1.00E-02	1.00E-02	0	0	0	0	0	0	0	0	0
Deinococcus_proteolyticus	0	0	0.15	0	0	0	0	0	0	0	0
Delftia_acidovorans	0	0	0	0	0	0	0	0	0	0	0
Desulfitobacterium_hafniense	0	0	0	0	0	0	0	0	0.14	0	0

Table 23. (cont.)

Desulfomonile_tiedjei	0	0	0	0	0	0	0	0	0	0	0
Dialister_invisus	0	0	0.75	0	0	0	0	0	1.1	0	0
Dichelobacter_nodosus	0	1.00E-02	0	0	0	0	0	0	0	0	0
Dickeya_dadantii	0	0	0	0	0	0	0	0	0	0	0
Dorea_longicatena	0	0	0	0	0	0	0	0	0	0	0
Enterobacter_asburiae	0	0	0	0	0	0	0	0	0	0	0
Enterobacter_cloacae	0	0	0	0	0	0	0	0	0	0	0
Enterobacter_sp._638	0	0	0	0	0	0	0	0	0	0	0
Enterococcus_faecalis	0	0	0	0	0	0	0	0	0	0	0
Enterococcus_faecium	1.00E-02	1.00E-02	0	0	0	0	0	0	0	0	0
Erysipelothrix_rhusiopathiae	0	0	0	0	0	0	0	1.00E-02	0	0	0
Erythrobacter_litoralis	0	0	0	0	0	0	0	0	0	0	0
Escherichia_coli	1.00E-02	1.00E-02	0.15	7.00E-02	3.00E-02	2.00E-02	2.00E-02	2.00E-02	0.14	3.00E-02	1.00E-02
Escherichia_fergusonii	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_eligens	0	0	0	0	0	0	0	4.00E-02	0	0	0
Eubacterium_hallii	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_limosum	0	0	0	0	0	0	0	0	0.14	0	0
Eubacterium_rectale	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_siraeum	0	0	0	0	0	0	0	0	0.14	0	0
Faecalibacterium_prausnitzii	0	0	0	0	0	4.00E-02	0	0	0	0	0
Ferroglobus_placidus	0	0	0	0	0	0	0	0	0	0	0
Fervidobacterium_pennivorans	0	0	0	0	0	0	0	0	0	0	0
Filifactor_alocis	0	0	0	0	0	0	0	0	0	0	0
Finegoldia_magna	0	0	1.35	0	0	0	0	0.66	4.66	0	1.00E-02
Flavobacterium_johnsoniae	1.00E-02	0	0	0	0	0	0	0	0	0	0
Fusobacterium_mortiferum	0	0	0	0	0	0	0	0	0	0	0
Fusobacterium_nucleatum	0	0	0.15	0	0	0	0	0	0.27	0	0
Fusobacterium_periodonticum	0	0	0	0	0	0	0	0	0	0	0
Gardnerella_vaginalis	70.7	66.94	0.9	92.38	92.33	0.96	5.58	43.95	0.82	3.00E-02	8.00E-02
Gemella_haemolysans	0	0	0	0	0	0	0	0	0	0	0
Gemella_morbillorum	0	0	0	0	0	0	0	0	0	0	0
Gemmatimonas_aurantiaca	0	0	0	0	0	0	0	0	0	0	0
Geobacter_uraniireducens	1.00E-02	0	0	0	0	0	0	0	0	0	0
Haemophilus_influenzae	0	0	0	0	0	0	0	0.13	0	0	0
Haemophilus_parainfluenzae	0	0	0	0	0	0	0	1.33	0	0	0
Halothermothrix_orenii	0	0	0	0	0	0	0	0	0.14	0	0
Helicobacter_bilis	2.00E-02	2.00E-02	0	0	0	0	0	0	0	0	0
Helicobacter_cetorum	0	0	0	0	0	0	0	0	0	0	0
Helicobacter_mustelae	0	0	0	0	0	0	0	0	0	0	0
Helicobacter_pylori	0	0	0	0	0	0	0	0	0.27	1.00E-02	2.00E-02
Hippea_maritima	0	1.00E-02	0	0	0	0	0	0	0	0	0
Holdemania_filiformis	0	0	0	0	0	0	0	0	0	0	0
Ignavibacterium_album	0	1.00E-02	0	0	0	0	0	0	0	0	0
Isoptericola_variabilis	0	0	0.15	0	0	0	0	0	0	0	0
Jannaschia_sp._ccs1	0	0	0	0	0	0	0	0	0	0	0
Jonquetella_anthropi	0	0	0	0	0	0	0	0	0	0	0
Kangiella_koreensis	0	0	0	0	0	0	0	0	0	0	0
Kineococcus_radiotolerans	0	0	0	0	0	0	0	0	0	0	0
Kitasatospora_setae	0	0	0	5.00E-02	5.00E-02	0	0	0	0	0	0
Klebsiella_oxytoca	0	0	0	0	0	0	0	0	0	0	0
Klebsiella_pneumoniae	0	0	0	0	0	0	0	0	0	0	0
Kribbella_flavida	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_acidophilus	0	0	0.15	0	5.00E-02	21.42	19.63	4.01	0.14	0.18	0.14
Lactobacillus_amylolyticus	0	0	0	0	0	1.02	0.83	0.28	0	0	0
Lactobacillus_amylovorus	0	0	0	4.00E-02	0	12.91	11.82	2.38	0.41	0	0
Lactobacillus_antri	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_casei	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_coleohominis	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_crispatus	0	0	0.3	2.00E-02	0	4.66	5.02	0.83	0.14	86.21	86.27
Lactobacillus_delbrueckii	0	0	0	0	0	0.2	0.27	4.00E-02	0	5.00E-02	2.00E-02
Lactobacillus_fermentum	0	0	0	0	0.24	0.76	2.37	4.00E-02	0	0	0
Lactobacillus_gasseri	19.73	22.7	0.3	0	0	6.00E-02	0.12	4.00E-02	0.27	0.18	8.00E-02
Lactobacillus_helveticus	1.00E-02	0	0	0	0	6.25	5.18	1.31	0	1.96	1.8
Lactobacillus_iners	2.00E-02	7.00E-02	0	0	0	0	0.15	0	0	8.00E-02	4.00E-02
Lactobacillus_jensenii	0	0	0	0	0	0.22	0.33	6.00E-02	0	2.00E-02	2.00E-02

Table 23. (cont.)

Lactobacillus_johnsonii	1.37	1.73	0	0	0	0.2	0.19	7.00E-02	0	0	0
Lactobacillus_kefiranofaciens	0	0	0	0	0	1.24	1.35	0.18	0	0	0
Lactobacillus_plantarum	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_reuteri	0	1.00E-02	0	0	0	0	0	0	0	1.00E-02	1.00E-02
Lactobacillus_rhamnosus	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_sakei	0	0	0	2.00E-02	0	0	0	0	0	0	0
Lactobacillus_ultunensis	0	0	0	0	0	2.7	2.6	0.44	0	0	0
Lactobacillus_vaginalis	0	0	0	0	0	0	0	0	0	0.31	0.4
Lactococcus_garvieae	0	0	0	0	0	0	0	0	0.14	0	0
Lactococcus_lactis	0	0	0	0	0	0	0	0	0	1.00E-02	0
Lawsonia_intracellularis	1.00E-02	1.00E-02	0.15	0	0	0	0	0	0	0	0
Leadbetterella_byssophila	0	0	0	4.00E-02	6.00E-02	0	0	0	0	0	0
Lysinibacillus_sphaericus	0	0	0	0	0	0	0	0	0	3.00E-02	5.00E-02
Mahella_australiensis	0	0	0	0	0	0	0	0	0.27	0	0
Mannheimia_haemolytica	0	0	0	0	0	0	0	1.00E-02	0	0	0
Marvinbryantia_formatexigens	0	0	0	0	0	0	0	1.00E-02	0.14	0	0
Mesoplasma_florum	0	0	0	0	0	0	0	0	0	0	0
Methanobrevibacter_ruminantium	1.00E-02	0	0	0	0	0	0	0	0	0	0
Methylacidiphilum_infernorum	3.00E-02	4.00E-02	0	0	0	0	0	0	0	0	0
Methylomicrobium_alcaliphilum	0	0	0	0	2.00E-02	0	0	1.00E-02	0	0	0
Micrococcus_luteus	0	0	0	0	0	0	0	0	0	0	0
Microlunatus_phosphovorus	0	0	0	0	0	0	0	1.00E-02	0	0	0
Micromonospora_sp._l5	0	0	0	0	0	0	0	1.00E-02	0	0	0
Mobiluncus_curtisii	0	0	2.84	0	0	0	0	0	0	0	0
Mycobacterium_abscessus	0	0	0	0	0	0	0	0	0	0	0
Mycobacterium_chubuense	0	0	0	0	0	0	0	0	0	0	0
Mycobacterium_smegmatis	0	0	0	0	0	0	0	0	0.14	0	0
Mycobacterium_vanbaalenii	0	0	0	0	0	0	0	0	0	0	0
Mycoplasma_arthritidis	0	0	0	0	0	0	0	0	0	0	0
Mycoplasma_crocodyli	0	0	0	0	0	0	0	0	0	0	0
Mycoplasma_gallisepticum	1.00E-02	0	0	0	0	0	0	1.00E-02	0	1.00E-02	4.00E-02
Mycoplasma_haemofelis	0	0	0	0	0	0	0	0	0	0	0
Mycoplasma_hominis	0	0	0	0	0	0	0	0	0	0	0
Mycoplasma_hyopneumoniae	1.00E-02	0	0	0	0	0.15	6.00E-02	1.00E-02	0.14	0	0
Mycoplasma_synoviae	0	0	0	0	0	0	0	0	0	0	1.00E-02
Nanoarchaeum_equitans	0	0	0	0	0	2.00E-02	0	0	0	0	0
Oenococcus_oeni	3.00E-02	4.00E-02	0	0	0	0	0	0	0	0	0
Opitutus_terrae	0	0	0.15	0	0	0	0	0	0	0	0
Orientia_tsutsugamushi	0	0	0	0	0	0	0	0	0	0	0
Oscillibacter_valericigenes	0	0	0	0	0	0	0	0	0	0	0
Owenweeksia_hongkongensis	0	1.00E-02	0	0	0	0	0	0	0	0	0
Pantoea_ananatis	0	0	0	0	0	0	0	0	0	0	0
Parabacteroides_distasonis	0	0	0	0	0	0	0	0	0	0	0
Parabacteroides_johnsonii	0	0	0	0	0	0	0	0	0	0	0
Parabacteroides_merdae	0	0	0	0	0	0	0	0	0	0	0
Parascardovia_denticolens	2.00E-02	5.00E-02	0	0	0	0	0	0	0	0	0
Parvimonas_micra	0	0	0.15	0	0	0	0	0	0	0	0
Pasteurella_multocida	0	0	0	0	0	0	0	0	0	0	0
Peptoniphilus_lacrimalis	0	0	0.15	0	0	0	0	0	0.82	0	0
Peptoniphilus_sp._oral_taxon_386	0	0	0.15	0	0	0	0	0	0	0	0
Peptostreptococcus_anaerobius	0	0	11.21	0	0	0	0	0	0.27	0	0
Peptostreptococcus_stomatis	0	0	0	0	0	0	0	1.00E-02	0	0	0
Phycisphaera_mikurensis	0	0	0	0	0	0	0	0	0	0	0
Polynucleobacter_necessarius	1.00E-02	1.00E-02	0	4.00E-02	5.00E-02	0	0	0	0	0	0
Porphyromonas_asaccharolytica	0	0	7.17	0	0	0	0	0	1.1	0	0
Porphyromonas_uenonis	0	1.00E-02	7.62	0	0	0	0	0	0.69	0	0
Prevotella_amnii	0	0	0	0	0	0	0	0	0	0	0
Prevotella_bergensis	0	0	0.9	0	0	0	0	0	0	0	0
Prevotella_bivia	0	0	0	0	0	0	0	0	0	0	0
Prevotella_buccalis	0	0	5.23	0	0	0	0	4.00E-02	0.27	0	0
Prevotella_copri	0	0	0	0	0	0	0	1.00E-02	0	0	0
Prevotella_melaninogenica	0	0	0	0	2.00E-02	0	0	0	0	0	0
Prevotella_ruminicola	0	0	0.15	0	0	0	0	0	0	0	0
Prevotella_sp._oral_taxon_299	0	0	0	0	0	0	0	0	0	0	0
Prevotella_timonensis	3.00E-02	6.00E-02	0.9	2.00E-02	0	0	0	7.00E-02	11.39	0	0

Table 23. (cont.)

Prochlorococcus_marinus	0	0	0	0	0	0	0	1.00E-02	0	2.00E-02	1.00E-02
Propionibacterium_acnes	0	0	0	0	0	0	0	0	0	1.00E-02	1.00E-02
Propionibacterium_freudenreichii	0	0	0.3	0	0	0	0	1.00E-02	0.14	0	0
Pseudoalteromonas_haloplanktis	0	0	0	0	0	0	0	0	0	0	0
Pseudomonas_aeruginosa	0	0	0	0	2.00E-02	0	0	0	0	0	0
Pseudomonas_fluorescens	0	0	0	0	0	0	0	0	0	0	0
Pseudomonas_mendocina	0	0	0	0	0	0	0	0	0	0	0
Pseudomonas_putida	0	0	0	0	0	0	0	1.00E-02	0	0	0
Pseudomonas_stutzeri	0	0	0	0	0	0	0	0	0	0	0
Psychromonas_ingrahamii	0	0	0	0	0	0	0	1.00E-02	0	0	0
Pyramidobacter_piscolens	0	0	0	0	0	0	0	0	0	0	0
Rhizobium_leguminosarum	0	0	0	0.11	2.00E-02	0	0	5.00E-02	0	0	0
Rhodococcus_equi	0	0	0	0	0	0	0	0	0	0	0
Rhodococcus_erythropolis	0	0	0	0	0	0	0	0	0	0	0
Rhodopseudomonas_palustris	0	0	0	0	0	0	0	0	0	0	0
Roseburia_hominis	0	0	0	0	0	0	0	0	0	0	0
Roseburia_intestinalis	0	0	0	0	0	0	0	0	0	0	0
Roseburia_inulinivorans	0	0	0	0	0	0	0	0	0	0	0
Rothia_mucilaginosa	0	0	0	0	0	0	0	0	0	0	0
Ruminococcus_albus	0	0	0	0	0	0	0	0	0.14	0	0
Ruminococcus_champanellensis	0	0	0	0	0	0	0	1.00E-02	0	0	0
Ruminococcus_lactaris	0	0	0	0	0	0	0	0	0	0	0
Ruminococcus_sp._sr1/5	0	0	0	0	0	0	0	0	0	0	0
Salinibacter_ruber	0	0	0.15	0	0	0	0	0	0	0	0
Salmonella_enterica	0	0	0	0	0	0	0	0	0	1.00E-02	0
Scardovia_inopinata	0	0	0	0	0	0	0	0	0	0	0
Selenomonas_ruminantium	0	0	0	0	0	2.00E-02	0	0	0	0	0
Serratia_symbiotica	0	0	0	0	0	0	0	0	0	0	0
Shewanella_baltica	0	0	0	0	0	0	0	0	0	0	0
Shigella_boydii	0	0	0.15	0	0	0	0	0	0	0	0
Shigella_dysenteriae	0	0	0	0	0	0	0	0	0	0	0
Shigella_flexneri	0	0	0	0	0	0	0	0	0	0	0
Shigella_sonnei	0	0	0	0	0	0	0	0	0	0	0
Shuttleworthia_satelles	0	0	0	0	0	0	0	0	0.14	0	0
Simkania_negevensis	0	0	0	0	0	0	0	0	0	0	0
Slackia_heliotrinireducens	0	0	0	0	0	0	0	0	0	0	0
Sorangium_cellulosum	0	0	0	0	0	0	0	0	0	0	0
Sphingomonas_wittichii	0	0	0	0	0	0	0	0	0	0	0
Sphingopyxis_alaskensis	0	0	0	0	0	0	0	0	0	0	0
Spirochaeta_africana	0	0	0	0	0	0	0	1.00E-02	0	0	0
Stackebrandtia_nassauensis	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_aureus	0	0	0	0	0	0	4.00E-02	0	0	0	0
Staphylococcus_epidermidis	0	0	0	0	0	0	0	0.15	0	2.00E-02	2.00E-02
Staphylococcus_haemolyticus	1.00E-02	0	0.15	0	0	0	0	0.21	0	0	0
Staphylococcus_hominis	0	0	0	2.00E-02	0	0	0	0.13	0	0	1.00E-02
Staphylococcus_lugdunensis	0	0	0	0	0	0	0	1.00E-02	0	0	0
Staphylococcus_pseudintermedius	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_saprophyticus	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_warneri	0	0	0	0	0	0	0	0	0	0	0
Streptobacillus_moniliformis	0	0	0	0	0	0	0	0	0.41	0	0
Streptococcus_agalactiae	0	0	0.15	2.00E-02	0	0	0	12.52	1.65	0	0
Streptococcus_anginosus	6.00E-02	2.00E-02	0	7.00E-02	2.00E-02	2.00E-02	2.00E-02	7.28	2.61	2.00E-02	0
Streptococcus_dysgalactiae	0	0	0	0	0	0	0	1.00E-02	0	0	0
Streptococcus_equi	0	0	0	0	0	0	0	4.00E-02	0.14	0	0
Streptococcus_equinus	0	0	0	0	0	0	0	4.00E-02	0.14	0	0
Streptococcus_gallolyticus	1.00E-02	0	0	0	0	0	0	0.12	0.27	0	0
Streptococcus_infantarius	0	0	0	0	0	0	0	5.00E-02	0	0	0
Streptococcus_intermedius	0	0	0	0	0	0	0	0.27	0.27	0	0
Streptococcus_mitis	0	0	0	0	0	0	0	8.00E-02	0	0	0
Streptococcus_mutans	0	0	0	0	0	0	0	5.00E-02	0.14	0	0
Streptococcus_oralis	6.00E-02	7.00E-02	0	0	0	0	0	7.00E-02	0	0	0
Streptococcus_parasanguinis	2.00E-02	0	0	0	0	0	0	0	0	0	0
Streptococcus_parauberis	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_pneumoniae	0	0	0	0	0	0	0	1.00E-02	0	0	0
Streptococcus_pseudopneumoniae	0	0	0	0	0	0	0	0	0	0	0

Table 23. (cont.)

Streptococcus_pyogenes	0	0	0	0	0	0	0	0.5	0.27	0	0
Streptococcus_salivarius	0	0	0	0	0	0	0	1.00E-02	0	0	0
Streptococcus_sanguinis	0	0	0	0	0	0	0	9.00E-02	0	0	0
Streptococcus_suis	0	0	0	0	0	0	6.00E-02	8.00E-02	0	3.00E-02	4.00E-02
Streptococcus_thermophilus	0	0	0	0	0	0	0	5.00E-02	0	0	0
Streptococcus_uberis	0	1.00E-02	0	0	0	0	0	0	0	0	0
Subdoligranulum_variabile	0	0	0	0	0	0	0	0	0	0	0
Sulfolobus_acidocaldarius	0	0	0	0	0	0	0	0	0	0	0
Sulfolobus_islandicus	0	0	0	0	0	0	0	0	0	0	0
Sulfolobus_solfataricus	0	0	0	0	0	0	0	0	0	0	0
Sulfolobus_tokodaii	1.00E-02	0	0	0	0	0	0	0	0	0	0
Sutterella_wadsworthensis	0	0	0	0	0	0	0	0	0	0	0
Synechocystis_sp._pcc_6803	0	0	0	0	0	0	0	0	0	0	0
Teredinibacter_turnerae	0	1.00E-02	0	0	0	0	0	0	0	0	0
Tetragenococcus_halophilus	0	0	0	0	0	0	0	0	0	0	0
Thermoanaerobacterium_thermosaccharolyticum	0	0	0	0	0	0	0	0	0	0	0
Thermobispora_bispora	0	0	0	2.00E-02	0	0	0	0	0	0	0
Thermosediminibacter_oceani	0	0	0	0	0	0	0	0	0.14	0	0
Thermovirga_lienii	0	1.00E-02	0	0	0	0	0	0	0.14	0	0
Turicella_otitidis	0	0	0	0	0	0	0	1.00E-02	0	0	0
Ureaplasma_parvum	0.34	0.43	0	0	0	0	0	0	0	0	0
Ureaplasma_urealyticum	0	3.00E-02	0	0	0	0	0	0	0	0	0
Variovorax_paradoxus	0	0	0	0	0	0	0	0	0	0	3.00E-02
Veillonella_parvula	0	0	0	0	0	0	0	0.49	0	0	0
Weeksella_virosa	0	0	0	0	0	0	0	0	0	0	0
Wolinella_succinogenes	0	0	0	0	0	0	0	0	0	0	0
Xanthomonas_campestris	0	1.00E-02	0.3	2.00E-02	0	0	0	0	0.82	0	0
Zymomonas_mobilis	0	0	0	0	0	0	0	0	0	3.00E-02	2.00E-02

Table 23. (cont.)

Bacterial species	205_3	205_5	205_6	206_1	206_2	206_3	206_4	206_5	206_6
[Bacteroides]_pectinophilus	0	0	0	0	0	0	0	0	0
[Clostridium]_bartlettii	0	0	0	0	0	0	0	0	0
[Clostridium]_sticklandii	0	0	0	0	0	0	0	0	0
[Ruminococcus]_gnavus	0	0	0	0	0	0	0	0	0
[Ruminococcus]_obeum	0	0	0	0	0	0	0	0	0
[Ruminococcus]_torques	0	0	0	0	0	0	0	0	0
Acetobacterium_woodii	0	0	0	0	0	0	0	0	0
Acholeplasma_laidlawii	0	0	0	0	0	0	0	0	0
Acidimicrobium_ferrooxidans	0	0	0	0	0	0	0	0	0
Acidothermus_cellulolyticus	0	0	0	0	1.00E-02	0	0	0	0
Aciduliprofundum_boonei	0	0	0	0	0	0	0	0	0
Acinetobacter_baumannii	0	0	0	0	0	0	0	0	0
Acinetobacter_junii	0	6.00E-02	0	0	0	0	0	0	0
Acinetobacter_lwoffii	0	6.00E-02	0	0	0	0	0	0	0
Actinobacillus_succinogenes	0	0	0	0	0	0	0	0	0
Actinomyces_coleocanis	0	0	0	0	0	0	0	0	0
Actinomyces_urogenitalis	0	6.00E-02	0	0	0	0	0	0	0
Aggregatibacter_aphrophilus	0	0	0	0	0	0	0	0	0
Alistipes_finegoldii	0	0	0	0	0	0	0	0	0
Alistipes_shahii	0	0	0	0	0	0	0	0	0
Alkaliphilus_metalliredigens	0	0	0	0	0	0	0	0	0
Allochromatium_vinosum	0	0	0	0	0	0	0	0	0
Alteromonas_macleodii	0	0	0	0	0	0	0	0	0
Alteromonas_sp._sn2	0	0	0	0	0	0	0	0	0
Amycolatopsis_mediterranei	0	0	0	0	0	0	0	0	0
Amycolicicoccus_subflavus	0	0	0	0	0	0	0	0	0
Anaerococcus_hydrogenalis	0	0	0	0	0	0	0	0	0.66
Anaerococcus_lactolyticus	0	0.36	0	0	0	0	0	0	1.76
Anaerococcus_prevotii	0	0	0	0	0	0	0	0	0.88
Anaerococcus_tetradius	0	0	0	0	0	0	0	0	0.22
Anaerococcus_vaginalis	0	0	0	0	0	0	0	0	5.51
Anaerotruncus_colihominis	0	0	0	0	0	0	0	0	0
Arcobacter_nitrofigilis	0	0	0	0	0	0	0	0	0
Atopobium_parvulum	0	6.00E-02	0	0	0	0	0	0	0
Atopobium_vaginae	0	0	0	0	0	0	0	0	0
Azospira_oryzae	0	0	0	0	0	0	0	0	0
Bacillus_cereus	0	0	0	0	0	4.00E-02	0	0	0
Bacillus_clausii	0	0	0	0	0	0	0	0	0
Bacillus_pseudofirmus	0	0	0	0	0	0	0	0	0
Bacillus_thuringiensis	0	0	0	0	0	0	0	0	0
Bacteriovorax_marinus	0	0	0	0	0	0	0	0	0
Bacteroides_caccae	0	0	0	0	0	0	0	0	0
Bacteroides_cellulosilyticus	0	0	0	0	0	0	0	0	0
Bacteroides_dorei	0	0	0	0	0	0	0	0	0
Bacteroides_fragilis	0	0	0	0	0	0	0	0	0
Bacteroides_ovatus	0	0	0	0	0	0	0	0	0
Bacteroides_stercoris	0	6.00E-02	0	0	0	0	0	0	0
Bacteroides_thetaiotaomicron	0	0	0	0	0	0	0	0	0
Bacteroides_uniformis	0	0	0	0	0	0	0	0	0
Bacteroides_vulgatus	0	6.00E-02	0	0	0	0	0	0	0
Bifidobacterium_adolescentis	0	0.12	0	0	0	0	0	0	0
Bifidobacterium_angulatum	0	0	0	0	0	0	0	0	0
Bifidobacterium_animalis	0	0	0	0	0	0	0	0	0
Bifidobacterium_bifidum	0	0.3	0	0	0	0	0	0	0
Bifidobacterium_breve	0	0	0	0	0	0	0	0	0
Bifidobacterium_catenulatum	0	6.00E-02	0	0	0	0	0	0	0
Bifidobacterium_dentium	0	0	0	5.00E-02	6.00E-02	0	0	4.00E-02	0
Bifidobacterium_gallicum	0	0	0	5.00E-02	4.00E-02	0	0	2.00E-02	0
Bifidobacterium_longum	0	6.00E-02	0	3.00E-02	1.00E-02	0	0	1.00E-02	0
Bifidobacterium_pseudocatenulatum	0	0	0	0	0	1.00E-02	0	2.00E-02	0.22
Bilophila_wadsworthia	0	0	0	0	0	0	0	0	0
Blautia_hydrogenotrophica	0	0	0	0	0	0	0	0	0
Brevibacterium_mcbrellneri	0	0	0	0	0	0	0	0	0.44

Table 23. (cont.)

Buchnera_aphidicola	0	0	0	0	0	0	0	0	0
Burkholderia_glumae	0	0	0	0	0	0	0	0	0
Burkholderia_phytofirmans	0	0	0	0	0	0	0	0	0
Burkholderia_pseudomallei	0	0	0	0	0	0	0	0	0
Butyrivibrio_crossotus	0	0	0	0	0	0	0	0	0
Butyrivibrio_fibrisolvens	0	0	0	0	0	0	0	0	0
Butyrivibrio_proteoclasticus	0	0	0	0	0	0	0	0	0.22
Caldicellulosiruptor_bescii	0	0	0	0	0	0	0	0	0
Campylobacter_fetus	0	0	0	0	0	0	0	0	0
Campylobacter_gracilis	0	0	0	0	0	0	0	0	0
Campylobacter_hominis	0	0	0	0	0	0	0	0	1.98
Campylobacter_lari	0	0	0	0	0	0	0	0	0
Campylobacter_upsaliensis	0	0	0	0	0	0	0	0	0
Candidatus_accumulibacter_phosphatis	0	0	0	0	0	0	0	0	0
Candidatus_azobacteroides_pseudotrichonymphae	0	0	0	0	0	1.00E-02	0.1	0	0
Candidatus_baumannia_cicadellinicola	3.00E-02	0	1.00E-02	0	0	0	0	0	0
Candidatus_carsonella_ruddii	0	0	0	0	0	0	0	0	0
Candidatus_nitrospira_defluvii	3.00E-02	0	2.00E-02	0	0	0	0	0	0
Candidatus_portiera_aleyrodidarum	0	0	0	0	0	0	0	0	0.22
Candidatus_puniceispirillum_marinum	0	6.00E-02	0	0	0	0	0	0	0
Candidatus_riesia_pediculicola	0	0	0	0	0	0	0	0	0
Candidatus_zinderia_insecticola	0	0	0	0	0	0	0	0	0
Carboxydothermus_hydrogenoformans	0	0	0	0	0	0	0	0	0.22
Catenibacterium_mitsuokai	0	0	0	0	0	0	0	0	0
Cellulophaga_algicola	0	0	0	0	0	0	0	0	0
Cellulosilyticum_lentocellum	0	0	0	6.00E-02	3.00E-02	0	0	2.00E-02	0
Chitinophaga_pinensis	0	0	0	0	0	0	0	0	0
Chlamydia_trachomatis	0	0	0	0	0	0	0	0	0
Chlamydophila_felis	0	0	0	0	0	0	0	0	0
Chryseobacterium_gleum	0	6.00E-02	0	0	0	0	0	0	0
Clostridiales_genomosp._bvab3	0	0	1.00E-02	0	0	0	0	0	0
Clostridium_acidurici	0	6.00E-02	0	0	0	0	0	0	0
Clostridium_botulinum	0	0	0	0	0	5.00E-02	0	0	0.44
Clostridium_cellulovorans	0	0	0	0	0	0	0	0	0
Clostridium_leptum	0	0	0	0	0	0	0	0	0
Clostridium_novyi	0	0	0	0	0	0	0	0	0
Clostridium_perfringens	0	0	0	0	0	0	0	0	0
Clostridium_sp._sy8519	0	0	0	0	0	0	0	0	0
Collinsella_aerofaciens	0	0.12	0	0	0	0	0	0	0
Comamonas_testosteroni	0	6.00E-02	0	0	0	0	0	0	0
Coprococcus_catus	0	0.12	0	0	0	0	0	0	0
Coprococcus_comes	0	0.12	0	0	0	0	0	0	0
Coprococcus_eutactus	0	0	0	0	0	0	0	0	0
Coprothermobacter_proteolyticus	0	0	0	0	0	0	0	0	0
Corynebacterium_accolens	0	6.00E-02	0	0	0	0	0	0	0
Corynebacterium_ammoniagenes	0	0.12	0	0	0	0	0	0	0
Corynebacterium_amycolatum	0	0.12	0	0	0	0	0	0	0.88
Corynebacterium_aurimucosum	0	0.65	0	0	0	0	0	0	0
Corynebacterium_glucuronolyticum	0	0.12	0	0	0	0	0	0	0
Corynebacterium_glutamicum	0	0	0	0	0	0	0	0	0
Corynebacterium_jeikeium	0	0	0	0	0	0	0	0	0
Corynebacterium_kroppenstedtii	0	0	0	0	0	0	0	0	0
Corynebacterium_lipophiloflavum	0	0.18	0	0	0	0	0	0	0
Corynebacterium_matruchotii	0	6.00E-02	0	0	0	0	0	0	0
Corynebacterium_pseudogenitalium	0	6.00E-02	0	0	0	0	0	0	0.22
Corynebacterium_striatum	0	0.18	0	0	0	0	0	0	0.44
Corynebacterium_tuberculostearicum	0	0.24	0	0	0	0	0	0	0.22
Corynebacterium_ulcerans	0	0	0	0	0	0	0	0	0
Corynebacterium_urealyticum	0	6.00E-02	0	0	0	0	0	0	0
Cronobacter_sakazakii	0	0	0	0	0	0	0	0	0
Cupriavidus_necator	0	0	0	0	0	0	0	0	0
Cytophaga_hutchinsonii	0	0	0	0	0	5.00E-02	0.1	0	0
Deinococcus_proteolyticus	0	0	0	0	0	0	0	0	0
Delftia_acidovorans	0	6.00E-02	0	0	0	0	0	0	0
Desulfitobacterium_hafniense	0	0	0	0	0	0	0	0	0

Table 23. (cont.)

Desulfomonile_tiedjei	0	0	0	0	0	0	0	0	0
Dialister_invisus	0	0	0	0	0	0	0	0	0.22
Dichelobacter_nodosus	0	0	0	0	0	0	0	0	0
Dickeya_dadantii	0	0	0	0	0	0	0	0	0
Dorea_longicatena	0	6.00E-02	0	0	0	0	0	0	0
Enterobacter_asburiae	0	0	0	0	0	0	0	0	0
Enterobacter_cloacae	0	0	0	0	0	0	0	0	0
Enterobacter_sp._638	0	0	0	0	0	0	0	0	0
Enterococcus_faecalis	0	0	0	0	0	0	0	0	0
Enterococcus_faecium	0	0	0	0	0	5.00E-02	0.1	0	0
Erysipelothrix_rhusiopathiae	0	0	0	0	0	0	0	0	0
Erythrobacter_litoralis	0	0.12	0	0	0	0	0	0	0
Escherichia_coli	1.00E-02	0	2.00E-02	2.00E-02	7.00E-02	4.00E-02	0.19	0	0.88
Escherichia_fergusonii	0	0	0	0	0	0	0	0	0
Eubacterium_eligens	0	6.00E-02	0	0	0	0	0	0	0
Eubacterium_hallii	0	0	0	0	0	0	0	0	0
Eubacterium_limosum	0	0	0	0	0	0	0	0	0
Eubacterium_rectale	0	6.00E-02	0	0	0	0	0	0	0
Eubacterium_siraeum	0	0	0	0	0	0	0	0	0
Faecalibacterium_prausnitzii	0	0.18	0	0	0	0	0	0	0
Ferroglobus_placidus	0	0	0	0	0	0	0	0	0
Fervidobacterium_pennivorans	0	0	0	0	0	0	0	0	0
Filifactor_alocis	0	0	0	0	0	0	0	0	0
Finegoldia_magna	0	0.42	0	0	0	0	0	0	1.54
Flavobacterium_johnsoniae	0	0	0	0	0	0	0	0	0
Fusobacterium_mortiferum	0	0	0	0	0	0	0	0	0
Fusobacterium_nucleatum	0	0.18	0	0	0	0	0	0	0
Fusobacterium_periodonticum	0	0	0	0	0	0	0	0	0
Gardnerella_vaginalis	5.00E-02	3.81	2.00E-02	91.51	90.45	1.2	36.67	85.16	29.3
Gemella_haemolysans	0	0	0	0	0	0	0	0	0
Gemella_morbillorum	0	0	0	0	0	0	0	0	0
Gemmatimonas_aurantiaca	0	0	0	0	0	0	0	0	0
Geobacter_uraniireducens	0	0	0	0	0	0	0	0	0
Haemophilus_influenzae	0	0	0	0	0	0	0	0	0
Haemophilus_parainfluenzae	0	6.00E-02	0	0	0	0	0	0	0
Halothermothrix_orenii	0	0	0	0	0	0	0	0	0
Helicobacter_bilis	0	0	0	0	1.00E-02	0	0	1.00E-02	0
Helicobacter_cetorum	0	0	0	0	0	1.00E-02	0	0	0
Helicobacter_mustelae	0	0	0	0	0	0	0	0	0
Helicobacter_pylori	2.00E-02	0	0	0	0	0	0	0	0
Hippea_maritima	0	0	0	0	0	0	0	0	0
Holdemania_filiformis	0	0	0	0	0	0	0	0	0.22
Ignavibacterium_album	0	0	0	0	0	0	0	0	0
Isoptericola_variabilis	0	0	0	0	0	0	0	0	0
Jannaschia_sp._ccs1	0	6.00E-02	0	0	0	0	0	0	0
Jonquetella_anthropi	0	0	0	0	0	0	0	0	0.66
Kangiella_koreensis	0	0	0	0	0	0	0	0	0
Kineococcus_radiotolerans	0	0	0	0	0	0	0	0	0
Kitasatospora_setae	0	0	0	0	0	0	0	0	0
Klebsiella_oxytoca	0	0	0	0	0	0	0	0	0
Klebsiella_pneumoniae	0	0	0	0	0	0	0	0	0
Kribbella_flavida	0	0	0	0	0	0	0	0	0
Lactobacillus_acidophilus	0.13	0.24	0.1	0	0	0	0	0	0
Lactobacillus_amylolyticus	0	0	1.00E-02	0	0	0	0	0	0
Lactobacillus_amylovorus	0	0	0	0	0	0	0	0	0
Lactobacillus_antri	0	0	0	0	0	0	0	0	0
Lactobacillus_casei	0	0	0	0	0	0	0	0	0
Lactobacillus_coleohominis	0	0	1.00E-02	0	0	0	0	0	0
Lactobacillus_crispatus	87.16	70.36	86.75	0.36	0.9	0.49	0.29	1.00E-02	0.22
Lactobacillus_delbrueckii	4.00E-02	0	3.00E-02	0	0	4.00E-02	0	0	0
Lactobacillus_fermentum	0	0	0	0	0	0	0	0	0
Lactobacillus_gasseri	0.22	0.54	1.00E-02	1.00E-02	0.2	83.89	49.61	2.45	1.98
Lactobacillus_helveticus	1.58	1.13	1.73	0	1.00E-02	2.00E-02	0	0	0
Lactobacillus_iners	8.00E-02	0.48	4.00E-02	0	0	0	1.26	0	2.2
Lactobacillus_jensenii	1.00E-02	0	3.00E-02	0	0	0	0	0	0

Table 23. (cont.)

Lactobacillus_johnsonii	1.00E-02	0	0	0	0	5.78	3.79	0.18	0.22
Lactobacillus_kefiranofaciens	0	0	0	0	0	0	0	0	0
Lactobacillus_plantarum	0	6.00E-02	0	0	0	0	0.1	0	0
Lactobacillus_reuteri	0	0	1.00E-02	0	0	0	0	0	0
Lactobacillus_rhamnosus	0	0	0	0.16	0.33	1.00E-02	0	0	0
Lactobacillus_sakei	0	0	0	0	0	0	0	0	0
Lactobacillus_ultunensis	0	0	0	0	0	0	0	0	0
Lactobacillus_vaginalis	0.46	0.42	0.2	0	0	1.00E-02	0	0	0
Lactococcus_garvieae	0	0	0	0	0	0	0	0	0
Lactococcus_lactis	1.00E-02	0	1.00E-02	0	0	0	0	0	0
Lawsonia_intracellularis	0	0	0	0	0	0	0	0	0
Leadbetterella_byssophila	0	0	0	0	0	0	0	0	0
Lysinibacillus_sphaericus	6.00E-02	6.00E-02	6.00E-02	0	1.00E-02	0	0	0	0
Mahella_australiensis	0	0	0	0	0	0	0	0	0
Mannheimia_haemolytica	0	0	0	0	0	0	0	0	0
Marvinbryantia_formatexigens	0	0	0	0	0	0	0	0	0
Mesoplasma_florum	0	0	0	0	0	0	0	0	0
Methanobrevibacter_ruminantium	0	0	0	0	0	1.00E-02	0	0	0
Methylacidiphilum_infernorum	0	0	0	0	0	0	0	0	0
Methylomicrobium_alcaliphilum	0	0	0	0	0	0	0	0	0
Micrococcus_luteus	0	0	0	0	0	0	0	0	0
Microlunatus_phosphovorus	0	0	0	0	0	0	0	0	0
Micromonospora_sp._l5	0	0	0	0	0	0	0	0	0
Mobiluncus_curtisii	0	0	0	0	0	0	0	0	0.88
Mycobacterium_abscessus	0	0	0	0	0	0	0	0	0
Mycobacterium_chubuense	0	6.00E-02	0	0	0	0	0	0	0
Mycobacterium_smegmatis	0	0.12	0	0	0	0	0	0	0
Mycobacterium_vanbaalenii	0	6.00E-02	0	0	0	0	0	0	0
Mycoplasma_arthritidis	0	0	0	0	0	0	0	0	0
Mycoplasma_crocodyli	0	0	0	0	0	0	0	0	0
Mycoplasma_gallisepticum	4.00E-02	0	1.00E-02	0	0	2.00E-02	0	0	0
Mycoplasma_haemofelis	0	0	0	0	0	0	0	0	0
Mycoplasma_hominis	0	0	0	0	0	0	0	0	0
Mycoplasma_hyopneumoniae	0	0	0	0	0	1.00E-02	0	0	0
Mycoplasma_synoviae	0	0	0	0	0	2.00E-02	0	0	0
Nanoarchaeum_equitans	0	0	0	0	0	0	0	0	0
Oenococcus_oeni	0	0	0	0	0	0	0	0	0
Opitutus_terrae	0	0	0	0	0	0	0	0	0
Orientia_tsutsugamushi	0	0	0	0	0	1.00E-02	0	0	0
Oscillibacter_valericigenes	0	0	0	0	0	0	0	0	0
Owenweeksia_hongkongensis	0	0	0	0	0	5.00E-02	0	0	0
Pantoea_ananatis	0	0	0	0	0	0	0	0	0
Parabacteroides_distasonis	0	0	0	0	0	0	0	0	0
Parabacteroides_johnsonii	0	0	0	0	0	0	0	0	0
Parabacteroides_merdae	0	0	0	0	0	0	0	0	0
Parascardovia_denticolens	0	0	0	9.00E-02	9.00E-02	0	0	7.00E-02	0
Parvimonas_micra	0	0	0	0	0	0	0	0	0
Pasteurella_multocida	0	0	0	0	0	0	0	0	0
Peptoniphilus_lacrimalis	0	0	0	0	0	0	0	0	0
Peptoniphilus_sp._oral_taxon_386	0	0	0	0	0	0	0	0	0
Peptostreptococcus_anaerobius	0	0.18	0	0	0	0	0	0	2.2
Peptostreptococcus_stomatis	0	0	0	0	0	0	0	0	0
Phycisphaera_mikurensis	0	0	0	1.00E-02	0	0	0	0	0
Polynucleobacter_necessarius	0	0	0	0	0	0	0	0	0
Porphyromonas_asaccharolytica	0	0	0	0	0	0	0	0	0.66
Porphyromonas_uenonis	0	0	0	0	0	0	0	0	1.32
Prevotella_amnii	0	0	0	0	0	0	0	0	0
Prevotella_bergensis	0	0	0	0	0	0	0	0	3.3
Prevotella_bivia	0	6.00E-02	0	0	0	0	0	1.14	0.22
Prevotella_buccalis	0	0	0	0	0	0	0	0	1.1
Prevotella_copri	0	0	0	0	0	0	0	0	0
Prevotella_melaninogenica	0	0	0	0	0	0	0	0	0
Prevotella_ruminicola	0	0	0	0	0	0	0	0	0
Prevotella_sp._oral_taxon_299	0	0	0	0	0	0	0	0	0
Prevotella_timonensis	0	0.36	0	0	0	0	0	0	20.26

Table 23. (cont.)

Prochlorococcus_marinus	0	0	3.00E-02	0	0	0	0	0	0
Propionibacterium_acnes	1.00E-02	0.36	1.00E-02	0	3.00E-02	0	0	1.00E-02	0.22
Propionibacterium_freudenreichii	0	0	0	0	0	0	0	0	0
Pseudoalteromonas_haloplanktis	0	0	0	0	0	0	0	0	0
Pseudomonas_aeruginosa	0	0	0	0	0	0	0	0	0
Pseudomonas_fluorescens	0	6.00E-02	0	0	0	0	0	0	0
Pseudomonas_mendocina	0	0	0	0	0	0	0	0	0
Pseudomonas_putida	0	0	0	0	0	0	0	0	0
Pseudomonas_stutzeri	0	0	0	0	0	0	0	0	0
Psychromonas_ingrahamii	0	0	0	0	0	0	0	0	0
Pyramidobacter_piscolens	0	0	0	0	0	0	0	0	0
Rhizobium_leguminosarum	0	0	0	0	0	0	0	0	0
Rhodococcus_equi	0	6.00E-02	0	0	0	0	0	0	0
Rhodococcus_erythropolis	0	0	0	0	0	0	0	0	0.22
Rhodopseudomonas_palustris	0	0	0	0	0	0	0	0	0
Roseburia_hominis	0	0	0	0	0	0	0	0	0
Roseburia_intestinalis	0	0	0	0	0	0	0	0	0
Roseburia_inulinivorans	0	0	0	0	0	0	0	0	0
Rothia_mucilaginosa	0	6.00E-02	0	0	0	0	0	0	0
Ruminococcus_albus	0	0	0	0	0	0	0	0	0
Ruminococcus_champanellensis	0	0	0	0	0	0	0	0	0
Ruminococcus_lactaris	0	0	0	0	0	0	0	0	0
Ruminococcus_sp._sr1/5	0	6.00E-02	0	0	0	0	0	0	0
Salinibacter_ruber	0	0	0	0	0	0	0	0	0
Salmonella_enterica	0	0	0	0	0	0	0	0	0
Scardovia_inopinata	0	0	0	0	0	0	0	0	0
Selenomonas_ruminantium	0	0	0	0	0	0	0	0	0
Serratia_symbiotica	0	0	0	0	0	0	0	0	0
Shewanella_baltica	0	0	0	0	0	0	0	0	0
Shigella_boydii	0	0	0	0	0	0	0	0	0
Shigella_dysenteriae	0	0	0	0	0	0	0	0	0
Shigella_flexneri	0	0	0	0	0	1.00E-02	0	0	0
Shigella_sonnei	0	0	0	0	0	0	0	0	0
Shuttleworthia_satelles	0	0	0	0	0	0	0	0	0
Simkania_negevensis	0	0	0	0	0	0	0	0	0
Slackia_heliotrinireducens	0	0	0	0	0	0	0	0	0
Sorangium_cellulosum	0	0	0	0	0	0	0	0	0
Sphingomonas_wittichii	0	0	0	0	0	0	0	0	0
Sphingopyxis_alaskensis	0	6.00E-02	0	0	0	0	0	0	0
Spirochaeta_africana	0	0	0	0	0	0	0	0	0
Stackebrandtia_nassauensis	0	6.00E-02	0	0	0	0	0	0	0
Staphylococcus_aureus	0	0	0	0	0	0	0	0	0
Staphylococcus_epidermidis	1.00E-02	0.42	1.00E-02	0	0	0	0	0	0
Staphylococcus_haemolyticus	0	0.18	0	0	0	0	0	0	0
Staphylococcus_hominis	0	0	0	0	0	0	0	0	0
Staphylococcus_lugdunensis	0	0	0	0	0	0	0	0	0
Staphylococcus_pseudintermedius	0	0	0	0	0	0	0	0	0
Staphylococcus_saprophyticus	0	0.12	0	0	0	0	0	0	0
Staphylococcus_warneri	0	6.00E-02	0	0	0	0	0	0	0
Streptobacillus_moniliformis	0	0	0	0	0	0	0	0	0
Streptococcus_agalactiae	0	0	0	0	0	0	0	1.00E-02	0
Streptococcus_anginosus	0	1.01	0	0	0	0	0	0	0
Streptococcus_dysgalactiae	0	0	0	0	0	0	0	0	0
Streptococcus_equi	0	0	0	0	0	0	0	0	0
Streptococcus_equinus	0	0	0	0	0	0	0	0	0
Streptococcus_gallolyticus	0	6.00E-02	0	0	0	0	0	0	0
Streptococcus_infantarius	0	0	0	0	0	0	0	0	0
Streptococcus_intermedius	0	0.12	0	0	0	0	0	0	0
Streptococcus_mitis	0	6.00E-02	0	0	0	0	0.19	1.38	1.1
Streptococcus_mutans	0	0	0	0	0	1.00E-02	0	0	0
Streptococcus_oralis	0	6.00E-02	0	0.14	4.00E-02	0	0	0.31	0.44
Streptococcus_parasanguinis	0	0	0	0	0	8.00E-02	0	1.00E-02	0
Streptococcus_parauberis	0	0	0	0	0	0	0	0	0
Streptococcus_pneumoniae	0	0	0	0	0	0	0.19	0.54	0.66
Streptococcus_pseudopneumoniae	0	6.00E-02	0	0	0	0	0	0.21	0.22

Table 23. (cont.)

Streptococcus_pyogenes	0	0	0	0	0	0	0	0	0
Streptococcus_salivarius	0	0	0	0	0	0	0	0	0
Streptococcus_sanguinis	0	0	0	0	0	1.00E-02	0	1.00E-02	0
Streptococcus_suis	2.00E-02	0	2.00E-02	0	0	0	0	0	0
Streptococcus_thermophilus	0	0	0	0	0	0	0	0	0
Streptococcus_uberis	0	0	0	0	0	0	0	0	0
Subdoligranulum_variabile	0	0.18	0	0	0	0	0	0	0
Sulfolobus_acidocaldarius	0	0	0	0	0	0	0	0	0
Sulfolobus_islandicus	0	0	0	0	0	0	0	0	0
Sulfolobus_solfataricus	0	0	0	0	0	0	0	0	0
Sulfolobus_tokodaii	0	0	0	0	0	0	0	0	0
Sutterella_wadsworthensis	0	0	0	0	0	0	0	0	0.44
Synechocystis_sp._pcc_6803	0	0	0	0	0	0	0	0	0
Teredinibacter_turnerae	0	6.00E-02	0	0	0	0	0	0	0
Tetragenococcus_halophilus	0	0	0	0	0	0	0	0	0
Thermoanaerobacterium_thermosaccharolyticum	0	0	0	0	0	0	0	0	0
Thermobispora_bispora	0	0	0	0	0	0	0	0	0
Thermosediminibacter_oceani	0	0	0	0	0	0	0	0	0
Thermovirga_lienii	0	0	0	2.00E-02	4.00E-02	0	0	2.00E-02	0
Turicella_otitidis	0	0	0	0	0	0	0	0	0
Ureaplasma_parvum	0	0	0	0	0	0	0	0	0
Ureaplasma_urealyticum	0	0	0	0	0	0	0	0	0
Variovorax_paradoxus	2.00E-02	0	1.00E-02	0	0	0	0	0	0
Veillonella_parvula	0	0	0	0	0	0	0	0	0
Weeksella_virosa	0	0	0	0	0	0	0	0	0
Wolinella_succinogenes	0	0	0	0	0	0	0	0	0
Xanthomonas_campestris	0	0	0	0	0	0	0	0	0
Zymomonas_mobilis	3.00E-02	0	2.00E-02	0	0	0	0	0	0

Note: X_X represents subject ID_time point. 1, 2, 3, 4, 5, and 6 represent samples collected at 8–12, 17–21, 26–30, and 35–38 weeks of gestation, during labor, and at 6 weeks post-partum, respectively.

Table 24. Taxonomic abundance determined from the metagenomic assembly in women with and without a previous preterm birth history (%).

Bacterial species	110_1	110_2	110_3	110_4	110_5	110_6	117_1	117_2	117_3	117_4	117_5
Acinetobacter_baumannii	0	0	0	0	0	0	0	0	0	0	0
Aerococcus_urinae	0	0	0	0	0	0	0	0	0	0	0
Aerococcus_viridans	0	0	0	0	0	0	0	0	0	0	0
Anaerococcus_hydrogenalis	0	0	0	0	0	0.03	0	0	0	0	0
Anaerococcus_lactolyticus	0	0	0.01	0	0	0.08	0	0	0	0	0
Anaerococcus_prevotii	0	0	0	0	0	0.02	0	0	0	0	0
Anaerococcus_tetradius	0	0	0	0	0	0.01	0	0	0	0	0
Atopobium_vaginae	0.06	0.17	0.41	0.04	0.24	4.54	0.39	0.03	0.09	0.08	0.11
Bacteroides_fragilis	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_adolescentis	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_breve	0.02	0.02	0.03	0.02	0.03	0.08	0.02	0.02	0.02	0.02	0.02
Bifidobacterium_dentium	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_longum	0	0	0.01	0	0.01	0.05	0	0	0	0	0
Bifidobacterium_thermophilum	0	0	0	0	0	0	0	0	0	0	0
Capnocytophaga_ochracea	0	0	0	0	0	0	0	0	0	0	0
Clostridiales_genomosp._BVAB3	0	0	0	0	0	0.01	0	0	0	0	0
Clostridium_pasteurianum	0	0	0	0	0	0	0	0	0	0	0
Clostridium_perfringens	0	0	0	0	0	0	0	0	0	0	0
Coprococcus_sp._ART55/1	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_lipophiloflavum	0	0	0	0	0	0	0	0	0	0	0
Dialister_microaerophilus	0	0	0	0	0	0	0	0	0	0	0
Eggerthella_sp._YY7918	0	0	0	0	0	0	0	0	0	0	0
Enterococcus_faecalis	0.1	0.22	0.52	0.05	0.25	6.8	0.26	0.03	0.1	0.08	0.18
Enterococcus_faecium	0.23	0.51	1.17	0.11	0.57	14.75	0.58	0.08	0.23	0.19	0.42
Eremococcus_coleocola	0	0	0	0	0	0	0	0	0	0	0
Escherichia_coli	0.03	0.05	0.17	0.02	0.14	1.3	0.08	0.01	0.03	0.03	0.04
Escherichia_fergusonii	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_siraeum	0	0	0	0	0	0	0	0	0	0	0
Finegoldia_magna	0	0	0.01	0	0.01	0.16	0	0	0	0	0
Frankia_symbiont_of_Datisca_glomerata	0	0	0	0	0	0	0	0	0	0	0
Gardnerella_vaginalis	0.26	0.17	0.27	0.06	0.2	1.44	0.11	0.06	0.07	0.07	0.09
Klebsiella_oxytoca	0	0	0	0	0	0.01	0	0	0	0	0
Klebsiella_pneumoniae	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_acidophilus	0	0	0.02	0	0.02	0.15	0.01	0	0	0	0
Lactobacillus_amylovorus	0	0.01	0.04	0.01	0.04	0.21	0.02	0	0.01	0.01	0.01
Lactobacillus_coleohominis	0	0	0	0	0	0.01	0	0	0	0	0
Lactobacillus_crispatus	0.53	0.49	1.66	98.67	96.33	2.21	92.56	98.82	98.41	98.28	96.46
Lactobacillus_delbrueckii	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_fermentum	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_gasseri	0.15	0.2	0.33	0.28	0.5	1.59	0.19	0.1	0.12	0.12	0.15
Lactobacillus_helveticus	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_iners	97.83	97.57	94.1	0.23	0.51	61.01	0.24	0.2	0.2	0.19	0.2
Lactobacillus_jensenii	0.19	0.2	0.41	0.33	0.58	1.59	5.11	0.47	0.47	0.73	2.06
Lactobacillus_johnsonii	0.22	0.16	0.16	0.06	0.06	0.15	0.06	0.07	0.06	0.05	0.06
Lactobacillus_kefiranofaciens	0	0	0.01	0	0.01	0.04	0.01	0	0	0	0
Lactobacillus_oris	0.08	0.06	0.06	0.03	0.03	0.06	0.03	0.03	0.03	0.02	0.03
Lactobacillus_reuteri	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_salivarius	0	0	0	0.01	0.01	0.01	0	0	0	0	0
Lactobacillus_vaginalis	0	0	0.02	0	0.02	0.12	0.03	0.03	0.04	0.02	0.04
Megasphaera_genomosp	0	0	0	0	0	0.01	0	0	0	0	0
Megasphaera_sp._UPII_135-E	0	0	0	0	0	0	0	0	0	0	0
Megasphaera_sp_UPII_199-6	0	0	0	0	0	0.01	0	0	0	0	0
Mobiluncus_curtisii	0.01	0.01	0.01	0	0	0.01	0	0	0	0	0
Mobiluncus_mulieris	0	0	0	0	0	0	0	0	0	0	0
Neisseria_gonorrhoeae	0.19	0.03	0.19	0.02	0.05	0.85	0.14	0.02	0.03	0.02	0.04
Odoribacter_splanchnicus	0	0	0	0	0	0	0	0	0	0	0
Peptoniphilus_duerdenii	0	0	0	0	0	0.04	0	0	0	0	0
Peptoniphilus_harei	0	0	0	0	0	0.03	0	0	0	0	0
Peptoniphilus_lacrimalis	0	0	0	0	0	0.05	0	0	0	0	0
Porphyromonas_asaccharolytica	0	0	0	0	0	0	0	0	0	0	0
Porphyromonas_gingivalis	0	0	0	0	0	0	0	0	0	0	0
Porphyromonas_uenonis	0	0	0	0	0	0.01	0	0	0	0	0
Prevotella_amnii	0	0	0	0	0	0.03	0	0	0	0	0

Table 24. (cont.)

Prevotella_bivia	0.02	0.03	0.13	0.02	0.13	0.75	0.05	0.01	0.02	0.02	0.02
Prevotella_buccalis	0	0	0	0	0	0.05	0	0	0	0	0
Prevotella_dentalis	0	0	0	0	0	0	0	0	0	0	0
Prevotella_denticola	0	0	0	0	0	0.05	0	0	0	0	0
Prevotella_disiens	0	0	0	0	0.01	0.08	0	0	0	0	0
Prevotella_intermedia	0	0	0	0	0	0	0	0	0	0	0
Prevotella_melaninogenica	0	0	0	0	0	0	0	0	0	0	0
Prevotella_oralis	0	0	0	0	0	0	0	0	0	0	0
Prevotella_timonensis	0	0	0	0	0	0.11	0	0	0	0	0
Proteus_mirabilis	0	0	0	0	0	0	0	0	0	0	0
Salmonella_enterica	0	0	0	0	0	0	0	0	0	0	0
Shigella_dysenteriae	0	0	0	0	0	0	0	0	0	0	0
Shigella_flexneri	0	0	0	0	0	0	0	0	0	0	0
Shigella_sonnei	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_aureus	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_lugdunensis	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_simulans	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_agalactiae	0.01	0.02	0.06	0.01	0.06	0.33	0.03	0	0.01	0.01	0.01
Streptococcus_bovis	0	0	0	0	0	0.07	0	0	0	0	0
Streptococcus_dysgalactiae	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_equi	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_gallolyticus	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_gordonii	0	0	0	0	0	0.01	0	0	0	0	0
Streptococcus_infantarius	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_intermedius	0	0	0	0	0	0.03	0	0	0	0	0
Streptococcus_macedonicus	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_mitis	0.01	0.01	0.04	0	0.04	0.3	0.02	0.01	0.01	0.01	0.01
Streptococcus_oligofermentans	0	0	0	0	0	0.01	0	0	0	0	0
Streptococcus_oralis	0	0	0	0	0	0.02	0	0	0	0	0
Streptococcus_parasanguinis	0	0	0	0	0	0.01	0	0	0	0	0
Streptococcus_pneumoniae	0.01	0.01	0.04	0.01	0.04	0.24	0.01	0	0.01	0.01	0.01
Streptococcus_pseudopneumoniae	0.01	0.01	0.03	0	0.03	0.16	0.01	0	0.01	0.01	0.01
Streptococcus_pseudoporcinus	0	0	0	0	0	0.02	0	0	0	0	0
Streptococcus_pyogenes	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_salivarius	0	0	0	0	0	0.01	0	0	0	0	0
Streptococcus_sanguinis	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_suis	0	0	0	0	0	0.01	0	0	0	0	0
Streptococcus_thermophilus	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_urinalis	0.03	0.03	0.03	0.01	0.02	0.11	0.01	0.01	0.01	0.01	0.01
Syntrophus_aciditrophicus	0	0	0	0	0	0	0	0	0	0	0
Tannerella_forsythia	0	0	0	0	0	0	0	0	0	0	0
Treponema_denticola	0	0	0	0	0	0	0	0	0	0	0
Ureaplasma_parvum	0	0	0.01	0	0.01	0.08	0	0	0	0	0
Ureaplasma_urealyticum	0	0	0	0	0	0.01	0	0	0	0	0
Veillonella_ratti	0	0	0	0	0	0.02	0	0	0	0	0
Vibrio_cholerae	0	0	0	0	0	0	0	0	0	0	0
Xanthomonas_campestris	0	0	0	0	0	0	0	0	0	0	0
Yersinia_pseudotuberculosis	0	0	0	0	0	0	0	0	0	0	0

Table 24. (cont.)

Bacterial species	117_6	119_1	119_2	119_3	119_4	119_6	120_1	120_2	120_3	120_4	120_6
Acinetobacter_baumannii	0	0	0	0	0	0	0	0	0	0	0.09
Aerococcus_urinae	0	0	0	0	0	0	0	0	0	0	0
Aerococcus_viridans	0	0	0	0	0	0	0	0	0	0	0
Anaerococcus_hydrogenalis	0	0	0	0	0	0.08	0	0	0	0	0
Anaerococcus_lactolyticus	0.01	0.01	0	0.01	0.01	0.35	0.02	0	0	0	0
Anaerococcus_prevotii	0	0	0	0	0	0.4	0.01	0	0	0	0
Anaerococcus_tetradius	0	0	0	0	0	0.3	0	0	0	0	0
Atopobium_vaginae	0.91	0.68	0.35	2.05	1.54	2.79	2.91	0.28	0.18	0.15	0.05
Bacteroides_fragilis	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_adolescentis	0.09	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_breve	54	0.04	0.22	0.08	0.02	37.38	0.3	0.09	0.08	0.02	0.05
Bifidobacterium_dentium	2.51	0	0.01	0	0	1.62	0	0	0	0	0.03
Bifidobacterium_longum	32.41	0.02	0.11	0.03	0.01	21.93	0.07	0.04	0.06	0	0.01
Bifidobacterium_thermophilum	0	0	0	0	0	0	0	0	0	0	0
Capnocytophaga_ochracea	0	0	0	0	0	0	0	0	0	0	0
Clostridiales_genomosp._BVAB3	0.02	0	0	0	0	0.05	0.06	0.08	0.15	0.01	0.02
Clostridium_pasteurianum	0	0	0	0	0	0	0	0	0	0	0
Clostridium_perfringens	0	0	0	0	0	0	0	0	0	0	0
Coprococcus_sp._ART55/1	0	0	0	0	0	0	0	0	0	0	0
Corynebacterium_lipophiloflavum	0	0	0	0	0	0	0	0	0	0	0.01
Dialister_microaerophilus	0	0	0	0	0	0.03	0	0	0	0	0
Eggerthella_sp._YY7918	0	0	0	0	0	0	0	0	0	0	0
Enterococcus_faecalis	0.88	1.04	0.31	1.69	1.51	4.44	5.18	0.57	0.27	0.22	0.14
Enterococcus_faecium	2.07	2.3	0.68	3.76	3.36	9.9	11.62	1.27	0.56	0.49	0.31
Eremococcus_coleocola	0	0	0	0	0	0	0	0	0	0	0
Escherichia_coli	0.49	0.28	0.09	0.49	0.35	0.91	1.31	0.08	0.14	0.11	92.15
Escherichia_fergusonii	0	0	0	0	0	0	0	0	0	0	0.02
Eubacterium_siraeum	0	0	0	0	0	0	0	0	0	0	0
Finegoldia_magna	0.02	0.01	0	0.02	0.01	0.44	0.05	0	0	0	0.01
Frankia_symbiont_of_Datisca_glomerata	0.03	0	0	0	0	0.02	0	0	0	0	0
Gardnerella_vaginalis	0.91	0.31	0.13	0.55	0.49	1.68	43.77	38.77	53.94	2.76	4.06
Klebsiella_oxytoca	0.01	0	0	0	0.01	0.01	0.02	0	0	0	1.78
Klebsiella_pneumoniae	0	0	0	0	0	0	0	0	0	0	0.04
Lactobacillus_acidophilus	0.06	0.02	0.01	0.06	0.04	0.09	0.09	0.01	0	0	0
Lactobacillus_amylovorus	0.11	0.05	0.02	0.1	0.07	0.15	0.21	0.01	0	0.01	0
Lactobacillus_coleohominis	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_crispatus	1.12	89.5	93.73	83.75	84.98	2.21	1.89	0.89	0.75	94.36	0.08
Lactobacillus_delbrueckii	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_fermentum	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_gasseri	0.8	0.48	0.24	0.84	0.61	1.19	25.32	56.52	42.59	0.59	0.15
Lactobacillus_helveticus	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_iners	0.65	0.44	0.24	0.8	0.77	1.61	1.79	0.34	0.36	0.2	0.06
Lactobacillus_jensenii	0.87	0.79	0.55	1.21	0.76	1.2	1.35	0.15	0.12	0.33	0.04
Lactobacillus_johnsonii	0.01	0.06	0.06	0.06	0.06	0.03	0.07	0.11	0.08	0.05	0
Lactobacillus_kefiranofaciens	0.03	2.73	2.83	2.67	2.52	0.05	0.04	0.01	0	0	0
Lactobacillus_oris	0.01	0.02	0.02	0.04	0.03	0.01	0.06	0.14	0.12	0.1	0
Lactobacillus_reuteri	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_salivarius	0	0.01	0.01	0.01	0.01	0.24	0	0	0	0	0
Lactobacillus_vaginalis	0.07	0.03	0.01	0.07	0.04	0.13	0.16	0.44	0.42	0.45	0.13
Megasphaera_genomosp	0.01	0	0	0	0	0.1	0	0	0	0	0
Megasphaera_sp._UPII_135-E	0	0	0	0	0	0.06	0	0	0	0	0
Megasphaera_sp_UPII_199-6	0	0	0	0	0	0.11	0	0	0	0	0
Mobiluncus_curtisii	0	0	0	0	0	0	0	0	0	0	0
Mobiluncus_mulieris	0	0	0	0	0	0	0	0	0	0	0
Neisseria_gonorrhoeae	0.27	0.25	0.14	0.54	0.51	1.37	1.22	0.11	0.05	0.06	0.02
Odoribacter_splanchnicus	0	0	0	0	0	0	0	0	0	0	0
Peptoniphilus_duerdenii	0.02	0.01	0	0.01	0.01	0.36	0.02	0	0	0	0
Peptoniphilus_harei	0	0	0	0	0	0.25	0.01	0	0	0	0
Peptoniphilus_lacrimalis	0.01	0	0	0.01	0.01	0.49	0.02	0	0	0	0
Porphyromonas_asaccharolytica	0	0	0	0	0	0	0	0	0	0	0
Porphyromonas_gingivalis	0	0	0	0	0	0	0	0	0	0	0
Porphyromonas_uenonis	0.01	0	0	0	0	0	0	0	0	0	0
Prevotella_amnii	0.02	0.01	0	0.01	0.01	0.05	0.02	0	0	0	0

Table 24. (cont.)

Prevotella_bivia	0.48	0.21	0.08	0.35	0.2	1.12	0.77	0.03	0.04	0.02	0.03
Prevotella_buccalis	0.06	0.01	0	0.01	0.01	0.06	0.03	0	0	0	0
Prevotella_dentalis	0	0	0	0	0	0	0	0	0	0	0
Prevotella_denticola	0.02	0	0	0.01	0	0.04	0.02	0	0	0	0
Prevotella_disiens	0.03	0.01	0	0.01	0	0.08	0.03	0	0	0	0
Prevotella_intermedia	0	0	0	0	0	0	0	0	0	0	0
Prevotella_melaninogenica	0	0	0	0	0	0	0	0	0	0	0
Prevotella_oralis	0	0	0	0	0	0	0	0	0	0	0
Prevotella_timonensis	0.04	0	0	0.01	0.01	0.09	0.04	0	0.01	0	0
Proteus_mirabilis	0	0	0	0	0	0	0	0	0	0	0.08
Salmonella_enterica	0	0	0	0	0	0	0	0	0	0	0.33
Shigella_dysenteriae	0	0	0	0	0	0	0	0	0	0	0
Shigella_flexneri	0	0	0	0	0	0	0	0	0	0	0.04
Shigella_sonnei	0	0	0	0	0	0	0	0	0	0	0.08
Staphylococcus_aureus	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_lugdunensis	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_simulans	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_agalactiae	0.22	0.37	0.04	0.18	1.43	0.44	0.32	0.01	0.01	0.01	0.01
Streptococcus_bovis	0.05	0.03	0	0.02	0.12	1.46	0.05	0	0	0	0.01
Streptococcus_dysgalactiae	0	0	0	0.01	0	0.02	0	0	0	0	0
Streptococcus_equi	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_gallolyticus	0	0	0	0	0	0.05	0	0	0	0	0
Streptococcus_gordonii	0	0	0	0	0	0.01	0.01	0	0	0	0
Streptococcus_infantarius	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_intermedius	0.01	0	0	0.01	0.03	1.09	0.01	0	0	0	0.01
Streptococcus_macedonicus	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_mitis	0.16	0.06	0.02	0.13	0.09	0.8	0.3	0.01	0.01	0.01	0.01
Streptococcus_oligofermentans	0.01	0	0	0	0	0.01	0	0	0	0	0
Streptococcus_oralis	0.01	0	0	0.01	0	0.09	0.02	0	0	0	0
Streptococcus_parasanguinis	0	0	0	0	0	0.07	0	0	0	0	0
Streptococcus_pneumoniae	0.15	0.06	0.02	0.15	0.08	0.72	0.31	0.01	0.01	0.01	0.01
Streptococcus_pseudopneumoniae	0.11	0.05	0.02	0.09	0.07	0.52	0.24	0.01	0	0.01	0.01
Streptococcus_pseudoporcinus	0.02	0.01	0	0.01	0.04	0.36	0.02	0	0	0	0
Streptococcus_pyogenes	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_salivarius	0	0	0	0	0	0.04	0	0	0	0	0
Streptococcus_sanguinis	0	0	0	0	0	0.03	0	0	0	0	0
Streptococcus_suis	0	0	0	0	0	0.05	0	0	0	0	0
Streptococcus_thermophilus	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_urinalis	0.11	0.04	0.01	0.04	0.12	0.69	0.05	0.01	0.01	0.01	0.01
Syntrophus_aciditrophicus	0	0	0	0	0	0	0	0	0	0	0.03
Tannerella_forsythia	0	0	0	0	0	0	0	0	0	0	0
Treponema_denticola	0	0	0	0	0	0.01	0	0	0	0	0
Ureaplasma_parvum	0.04	0.02	0	0.04	0.02	0.06	0.07	0	0	0	0
Ureaplasma_urealyticum	0	0	0	0	0	0	0	0	0	0	0
Veillonella_ratti	0.01	0	0	0.01	0	0.06	0.01	0	0	0	0
Vibrio_cholerae	0	0	0	0	0	0	0	0	0	0	0.02
Xanthomonas_campestris	0	0	0	0	0	0	0	0	0	0	0.02
Yersinia_pseudotuberculosis	0	0	0	0	0	0	0	0	0	0	0

Table 24. (cont.)

Bacterial species	125_1	125_2	125_3	125_4	125_6	201_1	201_2	201_3	201_5	202_1	202_2
Acinetobacter_baumannii	0	0	0	0	0	0	0	0	0	0	0
Aerococcus_urinae	0	0	0	0	0	0	0	0	0	0	0
Aerococcus_viridans	0	0	0	0	0	0	0	0	0	0	0
Anaerococcus_hydrogenalis	0	0	0	0	0	0	0	0	0	0	0
Anaerococcus_lactolyticus	0	0	0.01	0	0	0	0	0	0	0	0
Anaerococcus_prevotii	0	0	0	0	0	0	0	0	0	0	0
Anaerococcus_tetradius	0	0	0	0	0	0	0	0	0	0	0
Atopobium_vaginae	0.17	0.35	0.51	2.32	94.06	0.3	0.38	0.28	0.28	0.19	0.27
Bacteroides_fragilis	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_adolescentis	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_breve	0.01	0.01	0.02	0.04	0.08	0.02	0.02	0.02	0.02	0	0.13
Bifidobacterium_dentium	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_longum	0	0	0.01	0	0	0	0	0	0	0	0.01
Bifidobacterium_thermophilum	0	0	0	0.01	0.23	0	0	0	0	0	0
Capnocytophaga_ochracea	0	0	0	0	0	0	0	0	0	0	0
Clostridiales_genomosp._BVAB3	0.02	0.03	0.02	0.05	1.03	0	0	0	0	0	0.09
Clostridium_pasteurianum	0	0	0	0	0	0	0	0	0	0	0
Clostridium_perfringens	0	0	0	0	0	0	0	0	0	0	0
Coprococcus_sp._ART55/1	0	0	0	0	0.07	0	0	0	0	0	0.03
Corynebacterium_lipophiloflavum	0	0	0	0	0	0	0	0	0	0	0
Dialister_microaerophilus	0	0	0	0	0	0	0	0	0	0	0.12
Eggerthella_sp._YY7918	0	0	0	0	0.05	0	0	0	0	0	0
Enterococcus_faecalis	0.31	0.49	0.72	0.57	0.29	0.57	0.49	0.5	0.61	0	0.32
Enterococcus_faecium	0.68	1.09	1.65	1.38	0.64	1.18	0.98	0.98	1.24	0	0.72
Eremococcus_coleocola	0	0	0	0	0	0	0	0	0	0	0
Escherichia_coli	0.18	0.25	0.45	1.64	0.05	0.08	0.09	0.06	0.07	0.01	0.06
Escherichia_fergusonii	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_siraeum	0	0	0	0.02	0.74	0	0	0	0	0	0
Finegoldia_magna	0	0	0.01	0	0.01	0	0.01	0	0	0	0.09
Frankia_symbiont_of_Datisca_glomerata	0	0	0	0	0	0	0	0	0	0	0
Gardnerella_vaginalis	0.16	0.18	0.31	11.04	0.71	0.21	0.24	0.19	0.2	0.04	51.83
Klebsiella_oxytoca	0	0	0	0.03	0	0	0	0	0	0	0
Klebsiella_pneumoniae	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_acidophilus	0	0.01	0.02	0.02	0	0	0.01	0	0	0	0.01
Lactobacillus_amylovorus	0.01	0.01	0.03	0.02	0	0.01	0.02	0.01	0	0	0.01
Lactobacillus_coleohominis	0	0.01	0.01	0	0	0	0	0	0	0	0
Lactobacillus_crispatus	0.75	0.56	0.76	0.58	0.07	0.5	0.6	0.44	0.43	1.1	0.58
Lactobacillus_delbrueckii	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_fermentum	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_gasseri	17.78	0.33	0.51	0.81	0.09	0.27	0.26	0.28	0.29	94.36	43.16
Lactobacillus_helveticus	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_iners	52.21	60.91	53.34	55.08	0.52	95.34	95.26	95.82	95.54	0.26	0.57
Lactobacillus_jensenii	27.11	34.92	40.4	25.15	0.02	0.32	0.38	0.3	0.29	3.56	1.63
Lactobacillus_johnsonii	0.17	0.14	0.16	0.12	0.01	0.2	0.17	0.16	0.16	0.23	0.09
Lactobacillus_kefiranofaciens	0	0	0.01	0.01	0	0	0	0	0	0	0
Lactobacillus_oris	0.05	0.06	0.06	0.07	0	0.07	0.06	0.06	0.06	0.04	0.02
Lactobacillus_reuteri	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_salivarius	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_vaginalis	0.06	0.27	0.33	0.48	0.01	0.01	0.01	0.01	0.01	0	0.01
Megasphaera_genomosp	0.05	0.06	0.06	0.05	0.2	0	0	0	0	0	0
Megasphaera_sp._UPII_135-E	0	0	0	0	0	0	0	0	0	0	0.01
Megasphaera_sp_UPII_199-6	0.06	0.06	0.05	0.04	0.1	0	0	0	0	0	0
Mobiluncus_curtisii	0	0	0	0.01	0.04	0.01	0.01	0.01	0.01	0	0
Mobiluncus_mulieris	0	0	0	0	0	0	0	0	0	0	0
Neisseria_gonorrhoeae	0.11	0.08	0.17	0.13	0.05	0.22	0.19	0.18	0.13	0.06	0.07
Odoribacter_splanchnicus	0	0	0	0	0	0	0	0	0	0	0
Peptoniphilus_duerdenii	0	0	0.01	0	0	0	0	0	0	0	0
Peptoniphilus_harei	0	0	0	0	0	0.2	0.21	0.21	0.21	0	0
Peptoniphilus_lacrimalis	0	0	0	0	0.06	0	0	0	0	0	0.02
Porphyromonas_asaccharolytica	0	0	0	0	0	0	0	0	0	0	0
Porphyromonas_gingivalis	0	0	0	0	0.01	0	0	0	0	0	0
Porphyromonas_uenonis	0	0	0	0	0	0	0	0	0	0	0
Prevotella_amnii	0	0	0.01	0	0.03	0	0	0	0	0	0

Table 24. (cont.)

Prevotella_bivia	0.03	0.04	0.11	0.08	0.25	0.02	0.07	0.02	0.02	0.02	0.05
Prevotella_buccalis	0	0.01	0.02	0	0.08	0	0	0	0	0.01	0
Prevotella_dentalis	0	0	0	0	0	0	0	0	0	0	0
Prevotella_denticola	0	0	0	0	0.02	0	0	0	0	0	0
Prevotella_disiens	0	0	0.02	0	0.05	0	0	0	0	0.01	0
Prevotella_intermedia	0	0	0	0	0	0	0	0	0	0	0
Prevotella_melaninogenica	0	0	0	0	0	0	0	0	0	0	0
Prevotella_oralis	0	0	0	0	0	0	0	0	0	0	0
Prevotella_timonensis	0	0.01	0.02	0	0.09	0	0	0	0	0.01	0
Proteus_mirabilis	0	0	0	0	0	0	0	0	0	0	0
Salmonella_enterica	0	0	0	0.01	0	0	0	0	0	0	0
Shigella_dysenteriae	0	0	0	0	0	0	0	0	0	0	0
Shigella_flexneri	0	0	0	0	0	0	0	0	0	0	0
Shigella_sonnei	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_aureus	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_lugdunensis	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_simulans	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_agalactiae	0.01	0.02	0.05	0.04	0	0.01	0.03	0.01	0.01	0	0.02
Streptococcus_bovis	0	0	0.01	0	0	0.13	0.13	0.13	0.13	0.01	0
Streptococcus_dysgalactiae	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_equi	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_gallolyticus	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_gordonii	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_infantarius	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_intermedius	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_macedonicus	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_mitis	0.01	0.01	0.03	0.03	0	0.01	0.02	0.01	0	0	0.01
Streptococcus_oligofermentans	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_oralis	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_parasanguinis	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_pneumoniae	0.01	0.01	0.03	0.03	0	0.01	0.02	0.01	0.01	0	0.01
Streptococcus_pseudopneumoniae	0.01	0.01	0.02	0.02	0	0.01	0.02	0	0	0	0.01
Streptococcus_pseudoporcinus	0	0	0	0	0	0.26	0.24	0.25	0.25	0	0
Streptococcus_pyogenes	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_salivarius	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_sanguinis	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_suis	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_thermophilus	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_urinalis	0.02	0.02	0.03	0.03	0.14	0.03	0.03	0.03	0.03	0.02	0.01
Syntrophus_aciditrophicus	0	0	0	0	0	0	0	0	0	0	0
Tannerella_forsythia	0	0	0	0	0	0	0	0	0	0	0
Treponema_denticola	0	0	0	0	0	0	0	0	0	0	0
Ureaplasma_parvum	0	0.01	0.02	0.02	0	0	0.01	0	0	0.05	0.03
Ureaplasma_urealyticum	0	0	0	0	0	0	0	0	0	0	0
Veillonella_ratti	0	0	0	0	0.16	0	0	0	0	0	0
Vibrio_cholerae	0	0	0	0	0	0	0	0	0	0	0
Xanthomonas_campestris	0	0	0	0	0	0	0	0	0	0	0
Yersinia_pseudotuberculosis	0	0	0	0	0	0	0	0	0	0	0

Table 24. (cont.)

Bacterial species	202_3	202_4	202_6	203_1	203_2	203_3	203_4	203_5	203_6	205_1	205_2
Acinetobacter_baumannii	0	0	0	0	0	0	0	0.01	0	0	0
Aerococcus_urinae	0	0	0.01	0	0	0	0	0	0.05	0	0
Aerococcus_viridans	0	0	0	0	0	0	0	0	0	0	0
Anaerococcus_hydrogenalis	0	0	0.21	0	0	0	0	0.02	0.32	0	0
Anaerococcus_lactolyticus	0	0	0.93	0	0	0	0	0.12	1.78	0	0
Anaerococcus_prevotii	0	0	0.11	0	0	0	0	0.01	0.11	0	0
Anaerococcus_tetradius	0	0	0.18	0	0	0	0	0.01	0.19	0	0
Atopobium_vaginae	0.28	0.29	8.52	0.61	0.63	0.75	0.67	1.38	6.82	0.3	0.11
Bacteroides_fragilis	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_adolescentis	0	0	0	0	0	0	0	0	0	0	0
Bifidobacterium_breve	0.16	0.16	0.49	0.29	0.3	0.03	0.07	1.83	0.38	0.03	0.02
Bifidobacterium_dentium	0	0	0.01	0	0	0	0	0.08	0.02	0	0
Bifidobacterium_longum	0.01	0.01	0.07	0.03	0.04	0	0.02	0.98	0.21	0.01	0
Bifidobacterium_thermophilum	0	0	0	0	0	0	0	0	0	0	0
Capnocytophaga_ochracea	0	0	0.01	0	0	0	0	0	0.01	0	0
Clostridiales_genomosp._BVAB3	0.13	0.1	0.08	0	0	0	0	0.02	0.1	0	0
Clostridium_pasteurianum	0	0	0.03	0	0	0	0	0	0.01	0	0
Clostridium_perfringens	0	0	0.01	0	0	0	0	0	0	0	0
Coprococcus_sp._ART55/1	0.04	0.04	0.01	0	0	0	0	0	0.01	0	0
Corynebacterium_lipophiloflavum	0	0	0	0	0	0	0	0	0	0	0
Dialister_microaerophilus	0.16	0.14	0.05	0	0	0	0	0.01	0.05	0	0
Eggerthella_sp._YY7918	0	0	0	0	0	0	0	0	0	0	0
Enterococcus_faecalis	1.53	0.32	15.07	1.68	0.91	1.32	0.97	0.61	3.4	0.32	0.15
Enterococcus_faecium	3.77	0.71	33.3	3.74	2.01	3.05	2.31	1.34	7.39	0.69	0.35
Eremococcus_coleocola	0	0	0	0	0	0	0	0.02	0	0	0
Escherichia_coli	0.01	0.04	2.5	0.27	0.13	0.17	0.18	1.06	4.49	0.2	0.07
Escherichia_fergusonii	0	0	0	0	0	0	0	0	0	0	0
Eubacterium_siraeum	0	0	0.01	0	0	0	0	0	0.01	0	0
Finegoldia_magna	0.11	0.11	1.69	0.01	0	0.01	0.01	0.18	3.34	0.01	0
Frankia_symbiont_of_Datisca_glomerata	0	0	0	0	0	0	0	0	0	0	0
Gardnerella_vaginalis	65.43	63.81	3.45	90.93	94.45	0.84	5	47.48	6.47	0.23	0.16
Klebsiella_oxytoca	0	0	0.03	0	0	0	0	0.01	0.04	0	0
Klebsiella_pneumoniae	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_acidophilus	0	0	0.21	0.03	0.01	13.76	13.17	3.52	0.72	0.02	0.01
Lactobacillus_amylovorus	0	0	0.41	0.05	0.02	26.37	25.51	6.68	1.41	0.04	0.02
Lactobacillus_coleohominis	0	0	0.01	0	0	0	0	0	0.02	0	0
Lactobacillus_crispatus	0.29	0.39	2.73	0.28	0.12	42.88	41.32	12.52	10.12	96.04	97.69
Lactobacillus_delbrueckii	0	0	0	0	0	0.04	0.04	0.01	0.02	0	0
Lactobacillus_fermentum	0	0	0.01	0	0.02	0.05	0.11	0.01	0.01	0	0
Lactobacillus_gasseri	26.31	31.9	1.84	0.19	0.08	5.89	5.52	2.88	8.03	0.55	0.39
Lactobacillus_helveticus	0	0	0.01	0	0	0.77	0.81	0.21	0.04	0	0
Lactobacillus_iners	0.58	0.6	2.35	0.42	0.37	0.33	0.5	1.86	7.58	0.37	0.23
Lactobacillus_jensenii	0.89	1.03	1.77	0.22	0.07	0.76	0.84	2.33	9.68	0.53	0.38
Lactobacillus_johnsonii	0.06	0.08	0.07	0	0	1.4	1.34	0.37	0.06	0.05	0.05
Lactobacillus_kefiranofaciens	0	0	0.06	0.01	0	0.56	0.54	0.2	0.27	0.01	0
Lactobacillus_oris	0.01	0.02	0.01	0	0.03	0.09	0.15	0.02	0.04	0.03	0.03
Lactobacillus_reuteri	0	0	0	0	0	0	0	0	0	0	0
Lactobacillus_salivarius	0	0	0.01	0	0	0.07	0.1	0.09	0.02	0.01	0
Lactobacillus_vaginalis	0	0	0.2	0.02	0.01	0.02	0.04	0.18	0.75	0.09	0.11
Megasphaera_genomosp	0	0	0.13	0.06	0.07	0	0	0.07	0.21	0	0
Megasphaera_sp._UPII_135-E	0.01	0.01	0.91	0	0	0	0	0.02	0.48	0	0
Megasphaera_sp_UPII_199-6	0	0	0.15	0	0	0	0	0.03	0.28	0	0
Mobiluncus_curtisii	0.01	0.01	0	0	0	0	0	0	0.01	0	0
Mobiluncus_mulieris	0	0	0.01	0	0	0	0	0.01	0.01	0	0
Neisseria_gonorrhoeae	0.1	0.1	7.18	0.82	0.57	0.66	0.47	0.96	5.97	0.1	0.03
Odoribacter_splanchnicus	0	0	0.02	0	0	0	0	0	0.01	0	0
Peptoniphilus_duerdenii	0	0	1.26	0	0	0	0	0.1	1.36	0	0
Peptoniphilus_harei	0	0	0.33	0	0	0	0	0.06	0.36	0	0
Peptoniphilus_lacrimalis	0.03	0.03	0.74	0	0	0	0	0.16	0.85	0	0
Porphyromonas_asaccharolytica	0	0	0.37	0	0	0	0	0	0.03	0	0
Porphyromonas_gingivalis	0	0	0	0	0	0	0	0	0.01	0	0
Porphyromonas_uenonis	0	0	0.27	0	0	0	0	0	0.03	0	0
Prevotella_amnii	0	0	0.75	0	0	0	0	0.03	0.35	0	0

Table 24. (cont.)

Prevotella_bivia	0.03	0.03	2.83	0.1	0.03	0.05	0.1	0.99	5.03	0.12	0.06
Prevotella_buccalis	0	0	1.28	0	0	0	0	0.04	0.36	0	0
Prevotella_dentalis	0	0	0.03	0	0	0	0	0	0.01	0	0
Prevotella_denticola	0	0	0.8	0	0	0	0	0.03	0.31	0	0
Prevotella_disiens	0	0	2.25	0	0	0	0	0.05	0.6	0	0
Prevotella_intermedia	0	0	0.03	0	0	0	0	0	0.02	0	0
Prevotella_melaninogenica	0	0	0	0	0	0	0	0	0	0	0
Prevotella_oralis	0	0	0.04	0	0	0	0	0	0.01	0	0
Prevotella_timonensis	0	0	1.39	0	0	0	0	0.04	0.43	0	0
Proteus_mirabilis	0	0	0	0	0	0	0	0	0	0	0
Salmonella_enterica	0	0	0	0	0	0	0	0	0	0	0
Shigella_dysenteriae	0	0	0	0	0	0	0	0	0	0	0
Shigella_flexneri	0	0	0	0	0	0	0	0	0	0	0
Shigella_sonnei	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_aureus	0	0	0.01	0	0	0	0	0	0	0	0
Staphylococcus_lugdunensis	0	0	0	0	0	0	0	0.01	0.01	0	0
Staphylococcus_simulans	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_agalactiae	0	0.01	0.42	0.05	0.02	0.02	0.04	7.06	2.38	0.06	0.03
Streptococcus_bovis	0	0	0.18	0.01	0	0	0	0.76	0.3	0.01	0
Streptococcus_dysgalactiae	0	0	0.01	0	0	0	0	0.05	0.04	0	0
Streptococcus_equi	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_gallolyticus	0	0	0	0	0	0	0	0.03	0.02	0	0
Streptococcus_gordonii	0	0	0	0	0	0	0	0	0.01	0	0
Streptococcus_infantarius	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_intermedius	0	0	0.02	0	0	0	0	0.17	0.11	0	0
Streptococcus_macedonicus	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_mitis	0	0	0.23	0.03	0.01	0.02	0.03	0.52	1.43	0.04	0.02
Streptococcus_oligofermentans	0	0	0.01	0	0	0	0	0.02	0.05	0	0
Streptococcus_oralis	0	0	0.02	0	0	0	0	0.03	0.07	0	0
Streptococcus_parasanguinis	0	0	0.01	0	0	0	0	0.01	0.01	0	0
Streptococcus_pneumoniae	0	0	0.29	0.03	0.01	0.02	0.03	0.5	1.45	0.04	0.02
Streptococcus_pseudopneumoniae	0	0	0.23	0.02	0.01	0.01	0.02	0.35	1.04	0.03	0.01
Streptococcus_pseudoporcinus	0	0	0.46	0	0	0	0	0.27	0.53	0	0
Streptococcus_pyogenes	0	0	0	0	0	0	0	0.09	0.01	0	0
Streptococcus_salivarius	0	0	0.03	0	0	0	0	0.66	0.05	0	0
Streptococcus_sanguinis	0	0	0	0	0	0	0	0	0.01	0	0
Streptococcus_suis	0	0	0	0	0	0	0	0.02	0.03	0	0
Streptococcus_thermophilus	0	0	0	0	0	0	0	0	0	0	0
Streptococcus_urinalis	0.01	0.01	0.48	0.01	0.02	0.03	0.03	0.6	0.7	0.01	0.01
Syntrophus_aciditrophicus	0	0	0	0	0	0	0	0	0	0	0
Tannerella_forsythia	0	0	0	0	0	0	0	0	0	0	0
Treponema_denticola	0	0	0.01	0	0	0	0	0	0.04	0	0
Ureaplasma_parvum	0.02	0.04	0.04	0.01	0	0	0.01	0.1	0.45	0.01	0.01
Ureaplasma_urealyticum	0	0	0.01	0	0	0	0	0.01	0.03	0	0
Veillonella_ratti	0	0	0.35	0	0	0	0	0.04	0.48	0	0
Vibrio_cholerae	0	0	0	0	0	0	0	0	0	0	0
Xanthomonas_campestris	0	0	0	0	0	0	0	0	0	0	0
Yersinia_pseudotuberculosis	0	0	0	0	0	0	0	0	0	0	0

Table 24. (cont.)

Bacterial species	205_3	205_5	205_6	206_1	206_2	206_3	206_4	206_5	206_6
Acinetobacter_baumannii	0	0	0	0	0	0	0	0	0
Aerococcus_urinae	0	0	0	0	0	0	0	0	0
Aerococcus_viridans	0	0	0	0	0	0	0	0	0
Anaerococcus_hydrogenalis	0	0	0	0	0	0	0.01	0	0.04
Anaerococcus_lactolyticus	0	0.02	0	0	0	0	0.05	0	0.16
Anaerococcus_prevotii	0	0.01	0	0	0	0	0.01	0	0.03
Anaerococcus_tetradius	0	0.01	0	0	0	0	0.02	0	0.02
Atopobium_vaginae	0.11	2	0.06	0.43	0.83	0.28	3.24	0.02	6.83
Bacteroides_fragilis	0	0	0	0	0	0	0	0	0.01
Bifidobacterium_adolescentis	0	0	0	0	0	0	0	0	0
Bifidobacterium_breve	0.02	0.24	0.02	0.3	0.24	0.01	0.19	0.29	0.31
Bifidobacterium_dentium	0	0	0	0	0	0	0.01	0	0.01
Bifidobacterium_longum	0	0.06	0	0.01	0.01	0	0.09	0	0.11
Bifidobacterium_thermophilum	0	0	0	0	0	0	0	0	0
Capnocytophaga_ochracea	0	0	0	0	0	0	0	0	0
Clostridiales_genomosp._BVAB3	0	0.01	0	0.01	0.02	0	0.02	0.01	0.02
Clostridium_pasteurianum	0	0	0	0	0	0	0	0	0
Clostridium_perfringens	0	0	0	0	0	0	0	0	0
Coprococcus_sp._ART55/1	0	0	0	0	0	0	0	0	0
Corynebacterium_lipophiloflavum	0	0	0	0	0	0	0	0	0
Dialister_microaerophilus	0	0	0	0	0	0	0.01	0	0.02
Eggerthella_sp._YY7918	0	0	0	0	0	0	0	0	0
Enterococcus_faecalis	0.18	2.54	0.08	0.65	1.55	0.43	3.88	0.03	10.23
Enterococcus_faecium	0.41	5.61	0.17	1.47	3.55	0.95	8.51	0.06	22.73
Eremococcus_coleocola	0	0	0	0	0	0	0	0	0
Escherichia_coli	0.04	0.56	0.04	0.19	0.27	0.09	1.74	0.01	3.02
Escherichia_fergusonii	0	0	0	0	0	0	0	0	0
Eubacterium_siraeum	0	0	0	0	0	0	0.01	0	0.01
Finegoldia_magna	0	0.03	0	0.01	0.01	0	0.1	0	0.28
Frankia_symbiont_of_Datisca_glomerata	0	0	0	0	0	0	0	0	0
Gardnerella_vaginalis	0.09	2.85	0.07	94.05	89.48	1.16	19.06	87.64	19.3
Klebsiella_oxytoca	0	0.01	0	0	0	0	0.02	0	0.04
Klebsiella_pneumoniae	0	0	0	0	0	0	0	0	0
Lactobacillus_acidophilus	0.01	0.04	0.01	0.02	0.02	0.01	0.19	0	0.3
Lactobacillus_amylovorus	0.01	0.09	0.01	0.04	0.04	0.01	0.41	0	0.61
Lactobacillus_coleohominis	0	0	0	0	0	0	0.01	0	0.02
Lactobacillus_crispatus	97.72	80.59	98.51	1.04	1.8	1.83	3.94	0.15	4.64
Lactobacillus_delbrueckii	0	0	0	0	0	0	0	0	0.01
Lactobacillus_fermentum	0	0	0	0	0	0	0	0	0
Lactobacillus_gasseri	0.55	1.34	0.29	0.32	0.49	93.91	46.6	4.36	5.25
Lactobacillus_helveticus	0	0	0	0	0	0	0.01	0	0.01
Lactobacillus_iners	0.2	0.92	0.18	0.29	0.28	0.45	3.06	0.11	4.84
Lactobacillus_jensenii	0.32	0.78	0.33	0.28	0.31	0.26	3.13	0.02	4.26
Lactobacillus_johnsonii	0.05	0.05	0.06	0	0	0.21	0.14	0.01	0.03
Lactobacillus_kefiranofaciens	0	0.01	0	0.01	0.01	0	0.07	0	0.13
Lactobacillus_oris	0.04	0.03	0.03	0	0	0.08	0.05	0.01	0.02
Lactobacillus_reuteri	0	0	0	0	0	0	0	0	0
Lactobacillus_salivarius	0	0.01	0.01	0	0	0	0	0.02	0.01
Lactobacillus_vaginalis	0.12	0.1	0.05	0.02	0.03	0.01	0.24	0	0.4
Megasphaera_genomosp	0	0	0	0	0	0	0.01	0	0.02
Megasphaera_sp._UPII_135-E	0	0	0	0	0	0	0	0	0.01
Megasphaera_sp_UPII_199-6	0	0	0	0	0	0	0.01	0	0.04
Mobiluncus_curtisii	0	0	0	0	0	0	0	0	0
Mobiluncus_mulieris	0	0	0	0	0	0	0	0	0.01
Neisseria_gonorrhoeae	0.07	0.77	0.02	0.16	0.36	0.12	0.79	0.01	2.47
Odoribacter_splanchnicus	0	0	0	0	0	0	0	0	0.01
Peptoniphilus_duerdenii	0	0.02	0	0.01	0	0	0.05	0	0.16
Peptoniphilus_harei	0	0.01	0	0	0	0	0.01	0.01	0.08
Peptoniphilus_lacrimalis	0	0.01	0	0.26	0.26	0	0.07	0.18	0.18
Porphyromonas_asaccharolytica	0	0	0	0	0	0	0	0.01	0.04
Porphyromonas_gingivalis	0	0	0	0	0	0	0	0	0
Porphyromonas_uenonis	0	0	0	0	0	0	0	0.01	0.05
Prevotella_amnii	0	0.02	0	0	0	0	0.04	0.07	0.84

Table 24. (cont.)

Prevotella_bivia	0.02	0.37	0.02	0.14	0.13	0.05	1.39	2	3.95
Prevotella_buccalis	0	0.03	0	0	0	0	0.04	0.08	0.95
Prevotella_dentalis	0	0	0	0	0	0	0	0	0.01
Prevotella_denticola	0	0.01	0	0	0	0	0.03	0.06	0.51
Prevotella_disiens	0	0.04	0	0.01	0	0	0.04	0.13	1.42
Prevotella_intermedia	0	0.01	0	0	0	0	0	0.01	0.03
Prevotella_melaninogenica	0	0	0	0	0	0	0	0	0
Prevotella_oralis	0	0	0	0	0	0	0	0	0.03
Prevotella_timonensis	0	0.02	0	0	0	0	0.04	0.09	1.03
Proteus_mirabilis	0	0	0	0	0	0	0	0	0
Salmonella_enterica	0	0	0	0	0	0	0	0	0
Shigella_dysenteriae	0	0	0	0	0	0	0	0	0
Shigella_flexneri	0	0	0	0	0	0	0	0	0
Shigella_sonnei	0	0	0	0	0	0	0	0	0
Staphylococcus_aureus	0	0	0	0	0	0	0	0	0
Staphylococcus_lugdunensis	0	0	0	0	0	0	0	0	0
Staphylococcus_simulans	0	0	0	0	0	0	0	0	0
Streptococcus_agalactiae	0.01	0.13	0.01	0.07	0.08	0.02	0.73	0	1.01
Streptococcus_bovis	0	0.05	0	0.01	0	0	0.07	0.25	0.19
Streptococcus_dysgalactiae	0	0	0	0	0	0	0.01	0	0.02
Streptococcus_equi	0	0	0	0	0	0	0	0	0
Streptococcus_gallolyticus	0	0	0	0	0	0	0	0	0.01
Streptococcus_gordonii	0	0	0	0	0	0	0.01	0.03	0.01
Streptococcus_infantarius	0	0	0	0	0	0	0	0	0
Streptococcus_intermedius	0	0.01	0	0.01	0	0	0.02	0	0.04
Streptococcus_macedonicus	0	0	0	0	0	0	0	0	0
Streptococcus_mitis	0.01	0.15	0.01	0.05	0.05	0.02	0.54	1.65	0.99
Streptococcus_oligofermentans	0	0	0	0	0	0	0.02	0.05	0.01
Streptococcus_oralis	0	0.02	0	0	0	0	0.03	0.07	0.05
Streptococcus_parasanguinis	0	0	0	0	0	0	0	0	0
Streptococcus_pneumoniae	0.01	0.15	0.01	0.05	0.05	0.02	0.55	1.33	0.91
Streptococcus_pseudopneumoniae	0	0.12	0.01	0.03	0.03	0.01	0.37	1.12	0.61
Streptococcus_pseudoporcinus	0	0.02	0	0	0	0	0.02	0.03	0.11
Streptococcus_pyogenes	0	0	0	0	0	0	0	0	0.01
Streptococcus_salivarius	0	0	0	0	0	0	0.01	0.01	0.02
Streptococcus_sanguinis	0	0	0	0	0	0	0.01	0.01	0.01
Streptococcus_suis	0	0	0	0	0	0	0.01	0.01	0.01
Streptococcus_thermophilus	0	0	0	0	0	0	0	0	0
Streptococcus_urinalis	0.01	0.04	0.01	0.01	0.02	0.03	0.1	0.04	0.17
Syntrophus_aciditrophicus	0	0	0	0	0	0	0	0	0
Tannerella_forsythia	0	0	0	0	0	0	0	0	0
Treponema_denticola	0	0	0	0	0	0	0	0	0.01
Ureaplasma_parvum	0	0.02	0	0.01	0.02	0.01	0.13	0	0.17
Ureaplasma_urealyticum	0	0	0	0	0	0	0.01	0	0.02
Veillonella_ratti	0	0.01	0	0	0	0	0.01	0	0.09
Vibrio_cholerae	0	0	0	0	0	0	0	0	0
Xanthomonas_campestris	0	0	0	0	0	0	0	0	0
Yersinia_pseudotuberculosis	0	0	0	0	0	0	0	0	0

Note: X_X represents subject ID_time point. 1, 2, 3, 4, 5, and 6 represent samples collected at 8–12, 17–21, 26–30, and 35–38 weeks of gestation, during labor, and at 6 weeks post-partum, respectively.

CHAPTER 8. REFERENCES

1. Yeoman CJ, Thomas SM, Miller ME, et al. A multi-omic systems-based approach reveals metabolic markers of bacterial vaginosis and insight into the disease. *PLoS One*. 2013;8(2):e56111.

2. Savage DC. Microbial ecology of the gastrointestinal tract. *Annu Rev Microbiol*. 1977;31(6):107-133.

3. Braundmeier AG, Lenz KM, Chia N, et al. Individualized medicine and the microbiome of the reproductive tract. *Front Physiol*. 2015;6:97.

4. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Bacterial community variation in human body habitats across space and time. *Science*. 2009;326(5960):1694-1697.

5. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature*. 2012;486(7402):207-214.

6. Hickey RJ, Zhou X, Pierson JD, Ravel J, Forney LJ. Understanding vaginal microbiome complexity from an ecological perspective. *Transl Res*. 2012;160(4):267-282.

7. Ma B, Forney LJ, Ravel J. Vaginal microbiome: Rethinking health and disease. *Annu Rev Microbiol*. 2012;66:371-389.

8. Döderlein A. Das scheidensekret und seine bedeutung fur puerperalfieber. *Zentralblatt fur Bakteriology*. 1892;11:699.

9. Ravel J, Gajer P, Abdo Z, et al. Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A*. 2011;108 Suppl 1:4680-4687.

10. Romero R, Hassan SS, Gajer P, et al. The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. *Microbiome*. 2014;2(1):4-2618-2-4.

11. Farage M, Maibach H. Lifetime changes in the vulva and vagina. *Arch Gynecol Obstet*. 2006;273(4):195-202.

12. Gajer P, Brotman RM, Bai G, et al. Temporal dynamics of the human vaginal microbiota. *Sci Transl Med*. 2012;4(132):132ra52.

13. MacIntyre DA, Chandiramani M, Lee YS, et al. The vaginal microbiome during pregnancy and the postpartum period in a european population. *Scientific Reports*. 2015;5.

14. Walther-Antonio MR, Jeraldo P, Berg Miller ME, et al. Pregnancy's stronghold on the vaginal microbiome. *PLoS One*. 2014;9(6):e98514.

15. Aagaard K, Riehle K, Ma J, et al. A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. *PLoS ONE*. 2012;7:e36466.

16. Nott PN, Franklin M, Armitage C, Gelder MG. Hormonal changes and mood in the puerperium. *British Journal of Psychiatry*. 1976;128(4):379-383.

17. O'Hara MW, Schlechte JA, Lewis DA, Wright EJ. Prospective study of postpartum blues: Biologic and psychosocial factors. *Arch Gen Psychiatry*. 1991;48(9):801-806.

18. Boskey ER, Cone RA, Whaley KJ, Moench TR. Origins of vaginal acidity: High D/L lactate ratio is consistent with bacteria being the primary source. *Hum Reprod*. 2001;16(9):1809-1813.

19. Fettweis JM, Brooks JP, Serrano MG, et al. Differences in vaginal microbiome in african american women versus women of european ancestry. *Microbiology*. 2014;160(Pt 10):2272-2282.

20. Onderdonk AB, Delaney ML, Hinkson PL, Dubois AM. Quantitative and qualitative effects of douche preparations on vaginal microflora. *Obstet Gynecol*. 1992;80(3):333-338.

21. Klebanoff MA, Nansel TR, Brotman RM, et al. Personal hygienic behaviors and bacterial vaginosis. *Sex Transm Dis*. 2010;37(2):94-99.

22. Patterson JL, Stull-Lane A, Girerd PH, Jefferson KK. Analysis of adherence, biofilm formation and cytotoxicity suggests a greater virulence potential of gardnerella vaginalis relative to other bacterial-vaginosis-associated anaerobes. *Microbiology*. 2010;156(2):392-399.

23. Koumans EH, Sternberg M, Bruce C, et al. The prevalence of bacterial vaginosis in the united states, 2001-2004; associations with symptoms, sexual behaviors, and reproductive health. *Sex Transm Dis*. 2007;34(11):864-869.

24. Plitt SS, Garfein RS, Gaydos CA, Strathdee SA, Sherman SG, Taha TE. Prevalence and correlates of chlamydia trachomatis, neisseria gonorrhoeae, trichomonas vaginalis infections, and bacterial vaginosis among a cohort of young injection drug users in baltimore, maryland. *Sex Transm Dis*. 2005;32(7):446-453.

25. Schwebke JR. Gynecologic consequences of bacterial vaginosis. *Obstet Gynecol Clin North Am*. 2003;30(4):685-694.

26. Soper DE, Bump RC, Hurt WG. Bacterial vaginosis and trichomoniasis vaginitis are risk factors for cuff cellulitis after abdominal hysterectomy. *Am J Obstet Gynecol*. 1990;163(3):1016-21; discussion 1021-3.

27. Lamont RF, Sobel JD, Akins RA, et al. The vaginal microbiome: New information about genital tract flora using molecular based techniques. *Bjog-an International Journal of Obstetrics and Gynaecology*. 2011;118(5):533-549.

28. Sewankambo N, Gray RH, Wawer MJ, et al. HIV-1 infection associated with abnormal vaginal flora morphology and bacterial vaginosis. *Lancet*. 1997;350(9077):546-550.

29. Taha TE, Hoover DR, Dallabetta GA, et al. Bacterial vaginosis and disturbances of vaginal flora: Association with increased acquisition of HIV. *AIDS*. 1998;12(13):1699-1706.

30. Ralph S, Rutherford A, Wilson J. Influence of bacterial vaginosis on conception and miscarriage in the first trimester: Cohort study. *Br Med J*. 1999;319(7204):220-223.

31. LlahiCamp J, Rai R, Ison C, Regan L, TaylorRobinson D. Association of bacterial vaginosis with a history of second trimester miscarriage. *Human Reproduction*. 1996;11(7):1575-1578.

32. HAY P, LAMONT R, TAYLORROBINSON D, MORGAN D, ISON C, PEARSON J. Abnormal bacterial-colonization of the genital-tract and subsequent preterm delivery and late miscarriage. *Br Med J*. 1994;308(6924):295-298.

33. Larsson PG, Platz-Christensen JJ, Dalaker K, et al. 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*. 2000;79(5):390-396.

34. Watts DH, Krohn MA, Hillier SL, Eschenbach DA. Bacterial vaginosis as a risk factor for post-cesarean endometritis. *Obstet Gynecol*. 1990;75(1):52-58.

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

36. Azargoon A, Darvishzadeh S. Association of bacterial vaginosis, trichomonas vaginalis, and vaginal acidity with outcome of pregnancy. *Arch Iran Med*. 2006;9(3):213-217.

37. Hillier SL, Nugent RP, Eschenbach DA, et al. Association between bacterial vaginosis and preterm delivery of a low-birth-weight infant. *N Engl J Med*. 1995;333(26):1737-1742.

38. White BA, Creedon DJ, Nelson KE, Wilson BA. The vaginal microbiome in health and disease. *Trends in Endocrinology & Metabolism*. 2011;22(10):389-393.

39. Oleen-Burkey MA, Hillier SL. Pregnancy complications associated with bacterial vaginosis and their estimated costs. *Infect Dis Obstet Gynecol*. 1995;3(4):149-157.

40. Oduyebo OO, Anorlu RI, Ogunsola FT. The effects of antimicrobial therapy on bacterial vaginosis in non-pregnant women. *Cochrane Database Syst Rev*. 2009;(3):CD006055. doi(3):CD006055.

41. Hyman RW, Fukushima M, Diamond L, Kumm J, Giudice LC, Davis RW. Microbes on the human vaginal epithelium. *Proc Natl Acad Sci U S A*. 2005;102(22):7952-7957.

42. Hillier S. The vaginal microbial ecosystem and resistance to HIV. *AIDS Res Hum Retroviruses*. 1998;14:S17-S21.

43. Larsen B, Monif GR. Understanding the bacterial flora of the female genital tract. *Clin Infect Dis*. 2001;32:e69-e77.

44. Marrazzo JM, Koutsky LA, Eschenbach DA, Agnew K, Stine K, Hillier SL. Characterization of vaginal flora and bacterial vaginosis in women who have sex with women. *J Infect Dis*. 2002;185(9):1307-1313.

45. Redondo-Lopez V, Cook RL, Sobel JD. Emerging role of lactobacilli in the control and maintenance of the vaginal bacterial microflora. *Rev Infect Dis*. 1990;12:856-872.

46. Zhou X, Bent SJ, Schneider MG, Davis CC, Islam MR, Forney LJ. Characterization of vaginal microbial communities in adult healthy women using cultivation-independent methods. *Microbiology*. 2004;150(Pt 8):2565-2573.

47. Schwiertz A, Taras D, Rusch K, Rusch V. Throwing the dice for the diagnosis of vaginal complaints? *Ann Clin Microbiol Antimicrob*. 2006;5:4.

48. Africa CW, Nel J, Stemmet M. Anaerobes and bacterial vaginosis in pregnancy: Virulence factors contributing to vaginal colonisation. *Int J Environ Res Public Health*. 2014;11(7):6979-7000.

49. Spiegel CA. Bacterial vaginosis. *Clin Microbiol Rev*. 1991;4(4):485-502.

50. Fredricks DN, Fiedler TL, Marrazzo JM. Molecular identification of bacteria associated with bacterial vaginosis. *N Engl J Med*. 2005;353(18):1899-1911.

51. Srinivasan S, Fredricks DN. The human vaginal bacterial biota and bacterial vaginosis. *Interdiscip Perspect Infect Dis*. 2008;2008:750479.

52. Yeoman CJ, Yildirim S, Thomas SM, et al. Comparative genomics of gardnerella vaginalis strains reveals substantial differences in metabolic and virulence potential. *Plos One*. 2010;5(8):e12411.

53. Srinivasan S, Morgan MT, Fiedler TL, et al. Metabolic signatures of bacterial vaginosis. *MBio*. 2015;6(2):10.1128/mBio.00204-15.

54. AMSEL R, TOTTEN P, SPIEGEL C, CHEN K, ESCHENBACH D, HOLMES K. Nonspecific vaginitis - diagnostic-criteria and microbial and epidemiologic associations. *Am J Med*. 1983;74(1):14-22.

55. Chaijareenont K, Sirimai K, Boriboonhirunsarn D, Kiriwat O. Accuracy of nugent's score and each amsel's criteria in the diagnosis of bacterial vaginosis. *J Med Assoc Thai*. 2004;87(11):1270-1274.

56. NUGENT R, KROHN M, HILLIER S. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *J Clin Microbiol*. 1991;29(2):297-301.

57. Martinez I, Kim J, Duffy PR, Schlegel VL, Walter J. Resistant starches types 2 and 4 have differential effects on the composition of the fecal microbiota in human subjects. *PLoS One*. 2010;5(11):e15046.

58. Sha BE, Chen HY, Wang QJ, Zariffard MR, Cohen MH, Spear GT. Utility of amsel criteria, nugent score, and quantitative PCR for gardnerella vaginalis, mycoplasma

hominis, and lactobacillus spp. for diagnosis of bacterial vaginosis in human immunodeficiency virus-infected women. *J Clin Microbiol*. 2005;43(9):4607-4612.

59. Myziuk L, Romanowski B, Johnson SC. BVBlue test for diagnosis of bacterial vaginosis. *J Clin Microbiol*. 2003;41(5):1925-1928.

60. Wiggins R, Hicks SJ, Soothill PW, Millar MR, Corfield AP. Mucinases and sialidases: Their role in the pathogenesis of sexually transmitted infections in the female genital tract. *Sex Transm Infect*. 2001;77(6):402-408.

61. Briselden AM, Moncla BJ, Stevens CE, Hillier SL. Sialidases (neuraminidases) in bacterial vaginosis and bacterial vaginosis-associated microflora. *J Clin Microbiol*. 1992;30(3):663-666.

62. Cauci S, Driussi S, Monte R, Lanzafame P, Pitzus E, Quadrifoglio F. Immunoglobulin A response against gardnerella vaginalis hemolysin and sialidase activity in bacterial vaginosis. *Obstet Gynecol*. 1998;178(3):511-515.

63. Moncla BJ, Braham P, Hillier SL. Sialidase (neuraminidase) activity among gram-negative anaerobic and capnophilic bacteria. *J Clin Microbiol*. 1990;28(3):422-425.

64. Shah HN, Collins DM. Prevotella, a new genus to include bacteroides melaninogenicus and related species formerly classified in the genus bacteroides. *Int J Syst Bacteriol*. 1990;40(2):205-208.

65. Von Nicolai H, Hammann R, Salehnia S, Zilliken F. A newly discovered sialidase from gardnerella vaginalis. *Zentralblatt fur Bakteriologie Mikrobiologie und Hygiene - Abt.1 Orig.A*. 1984;258(1):20-26.

66. Monti E, Bonten E, D'Azzo A, et al. Sialidases in vertebrates: A family of enzymes tailored for several cell functions. *Advances in Carbohydrate Chemistry and Biochemistry, Vol 64*. 2010;64:403-479.

67. Corfield T. Bacterial sialidases - roles in pathogenicity and nutrition. *Glycobiology*. 1992;2(6):509-521.

68. Mcgregor JA, French JI, Jones W, et al. Bacterial vaginosis is associated with prematurity and vaginal fluid mucinase and sialidase - results of a controlled trial of topical clindamycin cream. *Obstet Gynecol*. 1994;170(4):1048-1060.

69. Olmsted SS, Meyn LA, Rohan LC, Hillier SL. Glycosidase and proteinase activity of anaerobic gram-negative bacteria isolated from women with bacterial vaginosis. *Sex Transm Dis*. 2003;30(3):257-261.

70. Wiggins R, Millar MR, Soothill PW, Hicks SJ, Corfield AP. Application of a novel human cervical mucin-based assay demonstrates the absence of increased mucinase activity in bacterial vaginosis. *Int J STD AIDS*. 2002;13(11):755-760.

71. Madhivanan P, Krupp K, Li T, et al. Performance of BVBlue rapid test in detecting bacterial vaginosis among women in mysore, india. *Infect Dis Obstet Gynecol*. 2014;2014:908313.

72. Bradshaw C, Morton A, Hocking J, et al. High recurrence rates of bacterial vaginosis over the course of 12 months after oral metronidazole therapy and factors associated with recurrence. *J Infect Dis*. 2006;193(11):1478-1486.

73. Ganu RS, Ma J, Aagaard KM. The role of microbial communities in parturition: Is there evidence of association with preterm birth and perinatal morbidity and mortality? *Am J Perinatol*. 2012.

74. Blencowe H, Cousens S, Oestergaard MZ, et al. 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. *The Lancet*. 2012;379(9832):2162-2172.

75. Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. *The Lancet*. 2008;371(9606):75-84.

76. DiGiulio DB, Romero R, Amogan HP, et al. Microbial prevalence, diversity and abundance in amniotic fluid during preterm labor: A molecular and culture-based investigation. *PLoS One*. 2008;3(8):e3056.

77. Zhou X, Brotman RM, Gajer P, et al. Recent advances in understanding the microbiology of the female reproductive tract and the causes of premature birth. *Infect Dis Obstet Gynecol*. 2010;2010:737425.

78. Gerber S, Vial Y, Hohlfeld P, Witkin SS. Detection of ureaplasma urealyticum in second-trimester amniotic fluid by polymerase chain reaction correlates with subsequent preterm labor and delivery. *J Infect Dis*. 2003;187(3):518-521.

79. Perni SC, Vardhana S, Korneeva I, et al. Mycoplasma hominis and ureaplasma urealyticum in midtrimester amniotic fluid: Association with amniotic fluid cytokine levels and pregnancy outcome. *Am J Obstet Gynecol*. 2004;191(4):1382-1386.

80. Han YW, Shen T, Chung P, Buhimschi IA, Buhimschi CS. Uncultivated bacteria as etiologic agents of intra-amniotic inflammation leading to preterm birth. *J Clin Microbiol*. 2009;47(1):38-47.

81. Agrawal V, Hirsch E. Intrauterine infection and preterm labor. *Semin Fetal Neonatal Med*. 2012;17(1):12-19.

82. Athayde N, Romero R, Maymon E, et al. Interleukin 16 in pregnancy, parturition, rupture of fetal membranes, and microbial invasion of the amniotic cavity. *Am J Obstet Gynecol*. 2000;182(1 Pt 1):135-141.

83. Wenstrom KD, Andrews WW, Hauth JC, Goldenberg RL, DuBard MB, Cliver SP. Elevated second-trimester amniotic fluid interleukin-6 levels predict preterm delivery. *Am J Obstet Gynecol*. 1998;178(3):546-550.

84. Kuczynski J, Lauber CL, Walters WA, et al. Experimental and analytical tools for studying the human microbiome. *Nat Rev Genet*. 2011;13(1):47-58.

85. Sharpton TJ. An introduction to the analysis of shotgun metagenomic data. *Front Plant Sci*. 2014;5:209.

86. Shah N, Tang H, Doak TG, Ye Y. Comparing bacterial communities inferred from 16S rRNA gene sequencing and shotgun metagenomics. *Pac Symp Biocomput*. 2011:165-176.

87. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: Proposal for the domains archaea, bacteria, and eucarya. *Proc Natl Acad Sci U S A*. 1990;87(12):4576-4579.

88. Hugenholtz P, Pace NR. Identifying microbial diversity in the natural environment: A molecular phylogenetic approach. *Trends Biotechnol*. 1996;14(6):190-197.

89. Hongoh Y, Yuzawa H, Ohkuma M, Kudo T. Evaluation of primers and PCR conditions for the analysis of 16S rRNA genes from a natural environment. *FEMS Microbiol Lett*. 2003;221(2):299-304.

90. Grice EA, Segre JA. The human microbiome: Our second genome. *Annu Rev Genom Hum Genet*. 2012;13:151-170.

91. Klappenbach JA, Saxman PR, Cole JR, Schmidt TM. Rrndb: The ribosomal RNA operon copy number database. *Nucleic Acids Res*. 2001;29(1):181-184.

92. Vetrovsky T, Baldrian P. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS One*. 2013;8(2):e57923.

93. Langille MG, Zaneveld J, Caporaso JG, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol*. 2013;31(9):814-821.

94. Yang F, Chia N, White BA, Schook LB. Compression-based distance (CBD): A simple, rapid, and accurate method for microbiota composition comparison. *BMC Bioinformatics*. 2013;14:136.

95. Kalyuzhnaya MG, Lapidus A, Ivanova N, et al. High-resolution metagenomics targets specific functional types in complex microbial communities. *Nat Biotechnol*. 2008;26(9):1029-1034.

96. Liu B, Gibbons T, Ghodsi M, Treangen T, Pop M. Accurate and fast estimation of taxonomic profiles from metagenomic shotgun sequences. *BMC Genomics*. 2011;12(SUPPL.2).

97. Mohammed MH, Ghosh TS, Reddy RM, Reddy CVSK, Singh NK, Mande SS. INDUS - A composition-based approach for rapid and accurate taxonomic classification of metagenomic sequences. *10th Int. Conference on Bioinformatics - 1st ISCB Asia Joint Conference 2011, InCoB 2011/ISCB-Asia 2011: Computational Biology - Proceedings from Asia Pacific Bioinformatics Network (APBioNet)*. 2011;12(SUPPL. 3).

98. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403-410.

99. McHardy AC, Martin HG, Tsirigos A, Hugenholtz P, Rigoutsos I. Accurate phylogenetic classification of variable-length DNA fragments. *Nat Methods*. 2007;4(1):63-72.

100. Nam-phuong Nguyen, Mirarab S, Liu B, Pop M, Warnow T. TIPP: Taxonomic identification and phylogenetic profiling. *Bioinformatics*. 2014;30(24):3548-3555.

101. Thomas T, Gilbert J, Meyer F. Metagenomics - a guide from sampling to data analysis. *Microbial informatics and experimentation*. 2012;2(1):3-3.

102. Imelfort M, Parks D, Woodcroft BJ, Dennis P, Hugenholtz P, Tyson GW. GroopM: An automated tool for the recovery of population genomes from related metagenomes. *PeerJ*. 2014;2:e603.

103. Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH. Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat Biotechnol*. 2013;31(6):533-538.

104. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015.

105. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2015;12(1):59-60.

106. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27-30.

107. Hendrick V, Altshuler LL, Suri R. Hormonal changes in the postpartum and implications for postpartum depression. *Psychosomatics*. 1998;39(2):93-101.

108. Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. *Appl Environ Microbiol*. 2008;74(8):2461-2470.

109. Yang F, Chia N, Mazur M, Pettigrew J, Schook LB, White BA. Genetically identical co-housed pigs as models for dietary studies of gut microbiomes. *Microbiome Sci. Med.* 2014;1(1):45-54.

110. Schachtschneider KM, Yeoman CJ, Isaacson RE, White BA, Schook LB, Pieters M. Modulation of systemic immune responses through commensal gastrointestinal microbiota. *PLoS One*. 2013;8(1):e53969.

111. Schloss PD, Westcott SL, Ryabin T, et al. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol*. 2009;75(23):7537-7541.

112. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naïve bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol*. 2007;73(16):5261-5267.

113. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. *Genome Biol*. 2011;12(6):R60-2011-12-6-r60.

114. Kruskal WILLIAMH, Wallis WALLEN. Use of ranks in one-criterion variance analysis. *Jour Amer Statist Assoc*. 1952;47((260)):583-621.

115. Goecks J, Nekrutenko A, Taylor J, Galaxy Team. Galaxy: A comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol*. 2010;11(8):R86.

116. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics*. 2014.

117. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nature Methods*. 2012;9(4):357-359.

118. Luo R, Liu B, Xie Y, et al. SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *Gigascience*. 2012;1(1):18-217X-1-18.

119. Assefa S, Keane TM, Otto TD, Newbold C, Berriman M. ABACAS: Algorithm-based automatic contiguation of assembled sequences. *Bioinformatics*. 2009;25(15):1968-1969.

120. Crusoe MR, Edvenson G, Fish J, et al. The khmer software package: Enabling efficient sequence analysis. . 2014.

121. Boisvert S, Raymond F, Godzaridis E, Laviolette F, Corbeil J. Ray meta: Scalable de novo metagenome assembly and profiling. *Genome Biol*. 2012;13(12):R122-2012-13-12-r122.

122. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079.

123. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*. 2010;11:119-2105-11-119.

124. Li W, Godzik A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*. 2006;22(13):1658-1659.

125. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760.

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

127. Ling Z, Kong J, Liu F, et al. Molecular analysis of the diversity of vaginal microbiota associated with bacterial vaginosis. *BMC Genomics*. 2010;11:488-2164-11-488.

128. Gelber SE, Aguilar JL, Lewis KLT, Ratner AJ. Functional and phylogenetic characterization of vaginolysin, the human-specific cytolysin from gardnerella vaginafis. *J Bacteriol*. 2008;190(11):3896-3903.

129. Kim TK, Thomas SM, Ho M, et al. Heterogeneity of vaginal microbial communities within individuals. *J Clin Microbiol*. 2009;47(4):1181-1189.

130. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics*. 2011;27(3):431-432.

131. Schweitzer L, Noblet J, Ruth E, Suffet IH, eds. *The formation, stability, and odor characterization of 2-ethyl-4-methyl-1,3-dioxolane (2-EMD).* ; 1999Water Science and Technology; No. 40.

132. Hay PE, Lamont RF, Taylor-Robinson D, Morgan DJ, Ison C, Pearson J. Abnormal bacterial colonisation of the genital tract and subsequent preterm delivery and late miscarriage. *BMJ*. 1994;308(6924):295-298.

133. Li J, McCormick J, Bocking A, Reid G. Importance of vaginal microbes in reproductive health. *Reproductive Sciences*. 2012;19(3):235-242.

134. Petrova MI, Lievens E, Malik S, Imholz N, Lebeer S. Lactobacillus species as biomarkers and agents that can promote various aspects of vaginal health. *Front Physiol*. 2015;6:81.

135. Martin R, Miquel S, Langella P, Bermudez-Humaran LG. The role of metagenomics in understanding the human microbiome in health and disease. *Virulence*. 2014;5(3):413-423.

136. Mirmonsef P, Hotton AL, Gilbert D, et al. Free glycogen in vaginal fluids is associated with lactobacillus colonization and low vaginal pH. *PLoS One*. 2014;9(7):e102467.

137. Hyman RW, Fukushima M, Jiang H, et al. Diversity of the vaginal microbiome correlates with preterm birth. *Reprod Sci*. 2014;21(1):32-40.

138. Romero R, Hassan SS, Gajer P, et al. The vaginal microbiota of pregnant women who subsequently have spontaneous preterm labor and delivery and those with a normal delivery at term. *Microbiome*. 2014;2:18-2618-2-18. eCollection 2014.

139. Lannon SM, Vanderhoeven JP, Eschenbach DA, Gravett MG, Adams Waldorf KM. Synergy and interactions among biological pathways leading to preterm premature rupture of membranes. *Reprod Sci*. 2014;21(10):1215-1227.

140. Koren O, Goodrich JK, Cullender TC, et al. Host remodeling of the gut microbiome and metabolic changes during pregnancy. *Cell*. 2012;150(3):470-480.