EVALUATING HISTORICAL PARADIGMS OF STERILITY IN PERINATAL MICROBIOLOGY AND RAMIFICATIONS FOR PREGNANCY OUTCOMES

by

JONATHAN MURRAY GREENBERG

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2020

MAJOR: BIOCHEMISTRY, MICROBIOLOGY, & IMMUNOLOGY

Approved by:

_____

Advisor Date

_____

_____

_____

_____

DEDICATION

This dissertation is dedicated to my parents and sister, who have all been immeasurably supportive over my graduate career and more importantly my entire life. I could not have done this without you.

ACKNOWLEDGMENTS

There is no way for me to express fully my sincere gratitude to all the individuals who have had an impact on my life and throughout my time here at Wayne State. None of it would have been possible without Dr. Kevin Theis. Rotating in his lab was one of the best decisions of my life and quite literally got me to where I am. His unwavering support throughout my time as a student and a scientist has meant the world to me, I am so grateful, and his guidance as a scientist and a professional has been fundamental to my development as a researcher. I would also like to thank the members of my committee: Dr. Withey, Dr. Pellett, Dr. Chen, and Dr. Luca who have supported me throughout my research, been patient, and given feedback that has undoubtedly contributed to my success.

From the Theis Lab, I must thank Dr. Andrew Winters, who has taught me many laboratory skills and helped sustain me by sharing many home-cooked meals. I also need to thank Madison Ahmad for always being willing to get coffee with me, supplying candy, and being a constant in the lab that I could always count on. I would also like to thank Emma Graffice, a new member of the lab, but one who quickly became invaluable and a supportive friend. I would like to thank Amruth Atyam, and the other undergraduate researchers who have spent time in the lab and allowed me to take on a mentorship role. And of course, I must thank the Perinatal Research Branch for allowing me to be a part of their work.

From the BMI department, I am grateful for the enormous support us students receive from April Wolak, Mary Dismuke, Joseph Fiore, and Lynette Ray. Mel Clay, you have contributed to all of our success by aiding us in the routine, but essential tasks that are required to do good science. Thank you to all the departmental faculty members that have played a role in making our experiences as students better through the various committees, and especially Drs. Holland and Thipparthi who were directly involved in my recruitment to Wayne State.

I have to thank my friends, new and old, and especially those I have met since moving to Detroit, I could not have been this successful without your support. And to Paul Breen, who has been a great friend and helped keep me accountable during our mutual tenure. And finally, thank you to my family, my parents Jeff and Liz Greenberg. Your love and encouragement have allowed me to come this far, and I cannot begin to fathom the sacrifices you have made in order to support me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Sterility has had a fundamental place in science and medicine since the introduction of the germ theory of disease, which posits that microscopic organisms cause contagious diseases [1-3]. The concept of sterility, as it applies to science and medicine, originated with the idea that an environment, whether it be a surface or a compartment of an organism or an inorganic object, is completely devoid of microorganisms; this includes protists, fungi, archaea, and bacteria [4, 5]. The value of sterility in the medical community became clear through the practice of surgery and the treatment of wounds [3, 6-8]. The importance of sterility developed into the practice of aseptic technique, which refers to the prevention of introducing foreign microorganisms into tissues or organs that do not normally contain these microorganisms [9, 10]. Aseptic technique entails the removal and eradication of any microorganisms from instruments and surfaces, thorough handwashing, and barrier protection of medical personnel directly involved in a procedure [4, 6]. The purpose of aseptic technique is to ensure the sterility of medical procedures and the safety of patients by protecting patients from the unintentional introduction of foreign and potentially infectious materials. The practice of aseptic techniques in medicine has saved millions of lives by preventing and reducing cases of infection and concomitantly increasing the success rate of medical procedures and surgeries [3, 6-8]. Similarly, in science, aseptic technique serves as a fundamental practice in conducting accurate, controlled, and replicable experiments that are un-confounded by unknown variables and contamination related to microorganisms [4, 9, 10]; this is necessary for the advancement and development of new technologies and practices.

Microbial communities, their diversity, and advancements in their characterization

Although sterile workspaces shall continue to be essential in the medical and scientific fields, microbial populations will undeniably continue to occupy and persist in nearly every exposed environment on the planet [11-13]. A microbial community is an ecological community made up of microorganisms [14, 15]. In a specific environment, it is often referred to as a microbiome, defined as the collective microorganisms, and the genetic and genomic potential of these microorganisms, inhabiting a particular environmental niche [16]. As sequencing technologies have improved and become more accessible over the past decade, microbiome research is no longer restricted to culture-dependent investigations and has resulted in a considerable increase in the attention and research on microbial communities. Still, much is unknown, including a great deal regarding the symbiotic relationships between microbiota (the members of a microbial community constituting a microbiome) and hosts. It is also important to note that while a microbial community and microbiome encompasses all microbial cells (bacteria, archaea, and microscopic eukaryotes), for all intents and purposes of this document, usage of the terms microbes, microorganisms, microbiome, and microbiota will be exclusively referring to bacteria.

Microorganisms are the most abundant life form on this planet and although their exact numbers are not possible to calculate, many estimates and formulations have made clear that we are living in a microbial world. For instance, our oceans alone contain more than 3×10^{28} bacteria, which vastly exceeds the number of stars in the visible universe [11, 12]. Not only are their numbers great, but microorganisms are capable of inhabiting the most extreme environments on the planet. Bacteria, for instance, have been detected in deep sea thermal vents [17], in the deepest layers of the Earth's crust [18], and in environments with extremely low [19] and high [20] pH environments – there is very little space on the planet that is unpopulated by

microorganisms. A recent estimate of the global diversity of bacteria and archaea suggests that there are more than a million species of bacteria on the planet [13]. The human body is no exception. Each human harbors more microorganisms in their gastrointestinal tract than there are people on the planet [21-23]. Indeed, all the external and mucosal surfaces of the human body are populated by diverse microbial communities. And yet, the historical presumption has been that the vast majority of internal organs of the human body are sterile.

Given that the human body consists of nearly equal numbers of human and bacterial cells (bacteria outnumber human cells if red blood cells are excluded), measuring around $1x10^{13}$ cells each [21, 24, 25], the importance of understanding the bacterial side of human health has been underappreciated and is gradually being rectified. Historically, microbial surveys in medical microbiology have been largely cultivation-based [26, 27], however, molecular surveys, especially those of phylogenetic marker genes (e.g. 16S rRNA gene) are becoming increasingly common [28-31]. Both methodological approaches have strengths and weaknesses, yet they can complement each other, thereby providing a more robust understanding of the microbial communities under investigation. Cultivating a microorganism from a clinical sample provides certainty of its viability [32, 33], and, once a microorganism is obtained in pure culture, its phenotype and genotype can be effectively characterized [34-36]. However, many microorganisms are recalcitrant to being cultivated in isolation within the laboratory [33, 36]. As such, cultivation-based surveys of microbial communities in clinical samples can preferentially select for community members that do grow well in the laboratory, thereby providing an incomplete or skewed representation of actual microbial community composition and structure [37-40]. With the development of next-generation sequencing (NGS) technologies, scientists can sequence millions of DNA molecules from clinical and research samples and describe the

genetic content in these samples to characterize their various biological features, including the compositions of microbial communities associated with these samples. More specifically, researchers now often utilize 16S rRNA gene sequencing to identify the compositions of microbial communities and metagenomic, or shotgun, sequencing to study the genomic and functional potential of these communities [41-53]. Culture-independent approaches like NGS surveys can provide more encompassing snapshots of microbial communities because they are not limited by the growth requirements of the bacteria [42, 43]. These new techniques facilitated the launch of the Human Microbiome Project, a large-scale NIH-funded collaborative project that utilized NGS in a major push towards understanding healthy human microbiota. Specifically, using 16S rRNA gene and metagenomic sequencing, the HMP characterized the bacterial communities of the human body including the mouth, skin, gut, and vagina [42]. The consortium of scientists and medical professionals involved in the HMP were tasked with investigating the microbiomes of various body sites of healthy individuals in order to establish a basic understanding of the diversity and function of the microorganisms in these environments. Understanding these communities in healthy individuals is a fundamental prerequisite to understanding their roles in disorders and diseases.

The human microbiome in the context of disease and low microbial biomass tissues

Since the inception of the HMP, thousands of studies have investigated various aspects of the human microbiome. Links to and associations with the human microbiome have been described in the context of multiple diseases and disorders, including *Clostridium difficile* infection (CDI) [54-56], obesity and Type II diabetes [57-59], dental caries [60-62], urinary tract disorders [48, 63-69], cystic fibrosis [70], inflammatory bowel disease [71-77], and asthma [78-80]. While the HMP and a majority of subsequent studies primarily focused on body sites known

to be inhabited by microorganisms, there are other body sites that have a much lower microbial load or have even been historically considered sterile. Next-generation sequencing technologies have allowed these body sites to be interrogated for potential microbial communities. Most descriptions of human-associated microbial communities have been from body sites with high microbial biomass, especially the gut [58, 59, 71, 81-96]. These early uses of NGS technologies to characterize the human microbiome had a high signal-to-noise ratio, meaning the samples being described had a high microbial DNA signal that far exceeded the inherent noise introduced through background DNA contamination in laboratory reagents and environments [97-99]. This high ratio of signal to noise results in the characterization of legitimate microbial signals in samples. However, as the microbial load in samples decreases, the signal-to-noise ratio decreases, or even inverts, such that that the background noise becomes a predominate portion of the data generated and is ultimately included within the characterization of the putative microbial community in a sample [99]. In these cases, most, if not all, legitimate biological signals can be overwhelmed and no longer distinguishable from contamination [97-104]. This caveat is one that must be addressed as researchers continue to investigate lower microbial biomass environments, including human body sites historically presumed to be sterile. As exquisitely demonstrated by Salter et. al. [99], simply by performing serial dilutions on cells from a pure culture of bacteria is sufficient to obscure legitimate signal. After five serial dilutions (e.g. 10^8 cells diluted to 10^3 cells), contaminating sequences from genera such as *Pseudomonas* and *Propionibacterium* comprised between 70-95% of the obtained microbial signal, with some variation depending upon the sequencing facility that was used to process the samples.

Molecular surveys failing to address caveats associated with low microbial biomass samples, specifically background DNA contamination, have led some to conclude the existence

of microbial communities in body sites historically considered sterile, such as the brain [105], blood [106, 107], endometrium [108, 109], placenta [44, 45, 49, 84, 95, 110-113], and bladder [114-118]. One significant and ubiquitous caveat with these investigations is that microbial genes are pervasive in the environment, even in the absence of microbial cells. This issue has led to much debate and premature conclusions could potentially influence medical science and even clinical care, which runs the risk of negatively impacting patient care and clinical outcomes. One of the principal ways this caveat can be properly addressed is through the inclusion of background DNA contamination controls (i.e. samples that receive no biological input and are processed and extracted alongside the biological samples) in all studies of low microbial biomass samples. These technical controls serve to account for the microbial DNA sequences that are ubiquitous in both DNA extraction kits and PCR reagents [98-100, 102, 104, 119]. Many studies have specifically evaluated the microbial signals from such technical controls and have described many microbial taxa as potential contaminants [98-100, 119]. For instance *Pseudomonas* [99], *Propionibacterium* [98, 99, 119], *Ralstonia* [98-100], *Corynebacterium* [99, 119], *Actinobacteria* [99, 119], *Burkholderia* [99, 100], *Escherichia* [102, 119], *Pelomonas* [98, 99, 119], and *Bradyrhizobium* [99, 100] are some of the more commonly reported taxa identified as contaminants in these studies. These microbial signals, which most often originate from DNA extraction kit reagents, have been given their own name, the “kitome” [98].

The kitome spreads uncertainty about the conclusions of microbiome work because it highlights the difficulty in knowing definitively if the presence of microbial DNA in a biological sample is legitimate or if it is a contaminant from the kitome. Several approaches and tools exist to help address this issue [120-122], however, there is no perfect way based solely on molecular surveys, which is why multiple methodologies are critical to validate DNA sequence-based

surveys. Various molecular techniques can be used to address the kitome such as targeted PCR, nested PCR, touchdown PCR, and metagenomic sequencing. While alternative PCR methods and metagenomics have their own biases (such as amplification biases that may exclude or underrepresent certain taxa [123-125] or limitations of taxonomic assignment based on reference databases) and can also be influenced by kitome contamination [98-100, 119], consistent trends in the data, such as one or two consistently detected and biologically relevant taxa across multiple investigatory approaches would help validate the potential microbiome data. For instance, conclusions from sequence-based surveys could be validated through Fluorescence *in situ* hybridization (FISH) [100], a molecular technique that allows for the visualization of bacterial cells within tissue samples; FISH is a critical tool for demonstrating the existence of a bacterial presence *in vivo*. While validating the presence of bacterial cells in tissue is important, cultivation persists as the most convincing evidence that bacterial cells are present in the biological samples and that they are alive. Again, both of these methodologies have their own caveats. In the case of FISH, fluorescent DNA probes can be cross-reactive or non-specific [126, 127]. For cultivation to be successful, knowledge of the appropriate and required growth conditions are necessary and many bacteria have unknown or unattainable culture conditions at this time [36], especially when dealing with mixed bacterial communities.

As previously mentioned, many studies investigating bacterial communities of these classically sterile sites have failed to employ multiple methodologies or failed to include appropriate kitome controls. This may lead to spurious conclusions, as these studies were missing critical information necessary for rejecting the null hypotheses that these sites are sterile. Bold claims suggesting a typical bacterial presence in tissues of the human body that are contrary

to historical views will certainly impact the practice of medicine and will, if incorrect, potentially result in improper patient care and treatment.

While most microbiome studies investigate body sites in the context of disease and/or rely on comparisons of a diseased cohort to a healthy cohort, this can be precarious when the organs under investigation are typically considered sterile. In one example, researchers compared 16S rRNA gene signals from post-mortem brains of Alzheimer's disease (AD) patients to those from brains of control patients [105]. This study makes two problematic assumptions: 1) there are bacteria in the brains of both cohorts, and 2) these bacteria were present before death and post-mortem removal of the organs. While the authors concluded that there exists an increase in bacteria within Alzheimer brain tissue over normal brain tissue, they failed to address the inherent caveats associated with low microbial biomass studies. First, although a significantly greater 16S rRNA gene sequence read count was found in the brains of AD patients compared to those from normal controls, sequence read count is not a quantitative measure appropriate for assessing bacterial load. The appropriate approach would be quantitative real-time PCR. Second, the taxonomic profiles indicated a high relative abundance of Actinobacteria in the brains of AD patients compared to controls, primarily *Propionibacterium acnes* (recently reclassified as *Cutibacterium acnes*). *P. acnes* is a ubiquitous human skin bacterium that is frequently detected in culture and molecular surveys as both a legitimate biological signal and as a contaminant [98, 99, 119, 128]. However, because this study was limited to a single methodology (i.e. 16S rRNA gene sequencing surveys), any conclusions about the existence of a brain microbiome should be considered premature. While 16S rRNA gene surveys alone may be appropriate for an exploratory study, to suggest paradigm-shifting conclusions like the brains of Alzheimer patients have more bacteria than normal patients is irresponsible and has the potential to cause harm.

Host-microbe interactions and consequences of premature shifts in sterility paradigms

The importance of accurately evaluating the presence of a microbial community and fully understanding its relationship with the human body has been made clear through a variety of noninfectious diseases that have been associated with microorganisms [43, 60, 65, 70, 75-77, 80, 108, 117, 129-133]. These associations reinforce the value of understanding where microbial communities begin and end in the human body. Should microbial communities be discovered in an anatomic environment that was previously believed to be sterile, it would open the door to new exploratory studies and research funding. However, if conclusions are drawn prematurely, it would put researchers on a fool's errand with expensive molecular surveys revealing the taxonomic and functional profiles of exogenous and contaminant DNA. More importantly would be the consequences on the medical community of the declaration that a body site presumed to be sterile is not actually sterile. If practitioners begin believing mistakenly that bacteria are normally present in a body site of healthy individuals, even at low or hard to detect numbers, practitioners may become laxer with safety and hygiene procedures. This would put patients at a high risk for infections or at the very least disruption and alteration of these communities. These theoretical consequences illustrate the diligence we must employ as researchers before presenting and advocating for shifts in classical paradigms of sterility.

Reevaluating paradigms of sterility in the context of pregnancy has been the focus of my dissertation research. Because infection is instrumental in pregnancy complications [134-146] and classically involves microorganisms reaching sterile tissues and causing inflammation, a thorough and complete understanding of the microbial signals in urogenital compartments is necessary for the best patient care and treatment. One important issue is that if there are actually resident microbiomes in the womb or bladder, it is likely that antibiotic use would affect these

microbial communities. While antibiotic therapy is generally only used when necessary, secondary consequences can result such as in the case of *Clostridium difficile* infection (CDI), which is a nosocomial infection that results from the use of antibiotics disturbing a patient's intestinal microbiota and allowing for *C. difficile* spores to colonize and populate the intestinal lining of a patient, causing chronic diarrhea and severe intestinal discomfort. Although antibiotics are often used during pregnancy, there may be consequences previously overlooked under assumptions of sterility in the bladder and womb. Also, under the presumption that these sites harbor microbiomes, new therapeutic and diagnostic tools, such as probiotics [69, 147, 148] or biomarker assays [122], can be explored. I must reiterate, however, these data and must be vetted before concluding that there are resident microbial communities in the bladder or womb.

CHAPTER 2: DOES THE HUMAN PLACENTA DELIVERED AT TERM HAVE A MICROBIOTA?

In the second chapter, I will present our study that addressed the literature asserting the existence of a placental microbiome. In 2014, Aagaard et al. published a pioneering study suggesting that a placental microbiome exists and that its primary member is *Escherichia coli* [44]. Since then, a wide range of studies from multiple groups have published reports either supporting Aagaard's claim [45, 49] or refuting it [128, 149, 150]. Chapter 2 will provide an overview of our own study that addressed the issues and oversights from the previous studies by approaching the hypothesis that there is a placental microbiome in term pregnancies by surveying cesarean delivered placentas with multiple methodologies and including extensive background DNA contamination controls. I will further describe and discuss the metagenomic analysis portion of the study that I performed. Additionally, I will describe a secondary analysis of the amplicon 16S rRNA gene sequence data wherein I address the same questions concerning sterility of the placenta in the context of two sequence classification methods, operational

taxonomic units (OTUs) [151] and amplicon sequence variants (ASVs) [152]. Operational taxonomic units were developed as a method for grouping 16S rRNA gene sequences based on a defined level of sequence similarity (e.g. 97 or 99%). Grouping sequences at a defined similarity eliminates some of the error inherent in sequencing technologies and permits a small degree of sequence variation while treating those sequences as though they came from the same microorganism. A key limitation of the OTU approach is that it lacks taxonomic resolution; the 16S rRNA gene sequences of multiple microorganisms, whether species or strains, can get grouped into the same OTU. This led to the proposal of ASVs as an alternative method for grouping and classifying sequences. This newer method increases the resolution of sequences to single nucleotide differences and this can be vital in differentiating different species and strains of bacteria. However, caveats exist with this method as well. By increasing the resolution to single nucleotide differences, researchers run the risk of characterizing sequences that are products of sequencing error and assuming that these sequences represent authentic and specific bacteria. Researchers are transitioning toward the use of ASVs as the dominant method for classifying 16S rRNA gene sequences, however, conclusions on which method is most appropriate need to be considered in various contexts, such as in investigations of low microbial biomass environments like the placenta.

CHAPTER 3: DOES THE MOUSE PLACENTA HAVE A MICROBIOTA? CULTURE AND MOLECULAR SURVEYS OF THE MURINE MICROBIOTA

The third chapter of my dissertation uses an animal model to more comprehensively address the sterility of the mammalian placental and fetal compartments. By transitioning from the human model to a mouse model we are able to address several questions that could not be answered using human participants. One benefit of using a mouse model is that sterile reproductive tissues and fetuses can be collected and investigated alongside other sterile and low

and high microbial biomass samples that can serve as controls for discriminating potential microbial signals. In this chapter, culture results from placental, fetal, and maternal tissues will be presented and analyzed extensively and compared to the microbial signals evident in murine tissues with high microbial load. Surprisingly few studies have assessed the broader mouse microbiome, outside of the intestine, and our data allow us to compare both molecular and culture-based surveys to determine if the murine placenta and fetus are sterile, compare the ability of culture to capture what is seen in 16S rRNA gene surveys, and to assess our results in the context of the established literature for culture and sequencing.

CHAPTER 4: A URINARY MICROBIOTA IN PREGNANCY: CULTIVATION- AND MOLECULAR-BASED COMPARISON OF FOLEY CATHETERIZED URINE, CLEAN CATCH URINE, AND VAGINAL SWABS FROM THE SAME WOMEN

The fourth chapter of my dissertation, while still investigating paradigms of sterility in perinatal medicine, shifts from the placenta to the female bladder in pregnancy. Like the placenta, the bladder has historically been considered a sterile organ. However, this is being reconsidered as recent molecular surveys of urine suggest the presence of a bladder microbiota [48, 63, 115-118, 153, 154]. Since urinary tract infections (UTIs) are the most frequent bacterial infection in women [155], and can lead to pregnancy complications such as spontaneous preterm birth [66, 134, 137, 138, 156-165], defining and characterizing bladder microbial communities could greatly alter how UTIs [69], and other urinary disorders [48, 63, 64, 67] are diagnosed and treated in pregnancy Additionally, urinary tract catheterization is standard practice for women delivering via cesarean section and such mechanical procedures can lead to UTIs [166] and potentially affect resident bladder microbial communities [65], Therefore, the objective of Chapter 4 is to evaluate the existence of a bladder microbiome in pregnant women, while also

assessing the effectiveness of different volume and collection methods for acquiring urine for microbiological investigations.

CHAPTER 5: GENERAL CONCLUSIONS

In closing, I provide a summary of the investigations I have completed, directions for future studies, and describe the significance of these findings in the context of low microbial biomass investigations.

CHAPTER 2: DOES THE HUMAN PLACENTA DELIVERED AT TERM HAVE A MICROBIOTA?

The metagenomic data presented in this chapter were a fundamental component of the manuscript "Does the human placenta delivered at term have a microbiota? Results of cultivation, quantitative real-time PCR, 16S rRNA gene sequencing, and metagenomics" published March 2019 in the *American Journal of Obstetrics and Gynecology* by Theis, K.R., Romero, R., Winters, A.D., **Greenberg, J.M.**, et al. [150]. Some of the text of this Chapter (i.e. Methods and Results) appears as it does in the published manuscript. I was an author on this manuscript and the data presented in this chapter include the portions of that manuscript with which I was most involved.

I. Abstract

Molecular surveys have allowed investigations of low microbial biomass tissues of the human body. The human placenta has been of particular interest, and its sterility has been the subject of recent debate. The majority of evidence suggesting there is a placental microbiota has relied on 16S rRNA gene sequencing and metagenomic sequencing without addressing the caveats inherent in low microbial biomass environments. The objective of this chapter was to determine whether human the placenta delivered at term has a microbiota using multiple modes of microbiologic inquiry and comparisons to background technical controls. Culture, quantitative real-time PCR (qPCR), and 16S rRNA gene sequencing were performed on 29 term cesarean placentas in addition to metagenomic sequencing on a subset of placentas. 28 out of 29 placentas were negative for culture and the one that was not had three unique bacterial colonies on a single plate that were not detected in 16S rRNA gene surveys of that placenta. Quantitative real-time PCR and 16S rRNA gene surveys indicated that the bacterial burden and profiles of placental samples were not distinct from technical controls. Metagenomic surveys of placental samples yielded limited bacterial signals and the predominant bacterial taxa included plant-associated and photosynthetic bacteria, which are ecologically implausible for internal tissues such as the placenta. Additional analyses of the 16S rRNA gene sequencing data, that utilized a new

classification method known as amplicon sequence variants (ASVs) and employed additional tools to control for background DNA contamination, supported the findings observed with the more traditional 16S rRNA gene classification method of operational taxonomic units (OTUs). Therefore, through multiple modes of microbiologic inquiry and computational approaches, we found a lack of evidence for a placental microbiota in the human placenta delivered at term.

II. Introduction

Most tissues within the human body are presumed to be sterile. In fact, nearly all internal sites, essentially anything not immediately adjacent to mucosal tissue, are viewed as being devoid of microbes. It is easier to list the mucosal or microbiota-containing sites of the body, i.e. nasal and oral cavities, lungs, gastrointestinal tract, vagina, and skin, than those considered sterile, i.e. most everything else. The presence of microbial communities in the urogenital and upper reproductive tracts, however, has been a source of much recent debate [44, 84, 108, 109, 133, 150, 167-169]. While for both sexes, the distal urethra indisputably contains microbes, the literature is less clear regarding sterility of the more proximal sites, up to and including the bladder (further discussed in Chapter 4). Additionally, while the microbial communities of the vagina and the cervix have been widely described in the literature [170-174], the existence of microbial communities in the endometrium of the uterus are far less conclusive [108, 109, 168]. In the context of pregnancy, because the placenta implants in the decidualized endometrium of the uterus, and because the placenta serves as the interfacing organ between the mother and fetus, the placenta has also become a recent target of microbiome investigation [44, 84, 113, 175].

It is well established that symbiotic microbes can colonize the human placenta and that this can negatively impact pregnancy outcomes [141, 176-182]. However, a unique, low

abundance, resident microbiota in the placentas of normal, asymptomatic pregnancies is now being proposed [44, 183]. Following a landmark study in 2012 by Aagaard et al., the placenta became a focus of microbiome work; this study made the bold claim that the placenta was not a sterile organ. The potential for the placenta to harbor a resident microbiota has implications not only for pregnancy outcome but also for the initial inoculation and development of fetal and early neonatal microbiota [183-186]. For instance, fetal inoculation with maternal microbes *in utero* via the placenta could serve to bolster and shape the early immune system and set the stage for the colonization and growth of site-specific (e.g., oral cavity, large intestine) microbial communities upon delivery [186-188]. However, in a recent comprehensive review of the subject, Perez-Munoz et al. [149] explained that current evidence in support of the *in utero* colonization hypothesis is unconvincing. The three principal concerns they highlighted are that supporting studies 1) have not demonstrated that the microbes captured via molecular surveys of the placenta and *in utero* environments were viable, 2) have not given sufficient consideration to the potential influences of mode of delivery on study outcomes, and they 3) lacked sufficient technical controls to address potential background contamination. Specifically, when characterizing and comparing the microbial profiles of low biomass samples via highly sensitive next-generation sequencing technologies, there is a substantial risk of amplifying, sequencing, and consequently describing contaminating DNA that is unavoidably present in extraction kits, PCR and sequencing reagents, and broader laboratory environments [98, 99, 119, 128, 149, 189]. The high sensitivity afforded by the technology also provides susceptibility to false positives. Notably, one preliminary study that did incorporate sufficient technical controls for background DNA contamination found similar concentrations of bacterial DNA in placental tissue and control samples, and the bacterial profiles of the two were not distinguishable [128]. Lauder et al.

[128] concluded that, when working with low microbial biomass samples and capitalizing on next-generation sequencing technologies, begin with the null hypothesis that microbial signals obtained from the biological samples are contamination only. Perez-Munoz et al. [149] additionally proposed that there is a particular need for studies evaluating this null hypothesis in healthy term pregnancies, with delivery via caesarean section, in which corroborative evidence is evaluated across multiple modes of microbiologic inquiry. Therefore, the primary objective of the study featured in this chapter was to evaluate the existence of a unique resident microbiota in the placenta from normal term pregnancies using multiple modes of microbiologic inquiry and incorporating rigorous technical controls for DNA contamination.

As described in Theis et al. [150], the placentas of 29 women delivering via elective cesarean section at term were evaluated for evidence of a microbiota using bacterial culture, 16S rRNA gene quantitative real-time PCR, 16S rRNA gene sequencing, and metagenomics. Culture results were negative for all but one placenta, from which three unique colonies were recovered from a single agar growth plate. These microbes were common environmental bacteria and their 16S rRNA genes were not detected in the 16S rRNA gene surveys of the placental sample from which they were "recovered," suggesting that they were likely laboratory contaminants. Both 16S rRNA gene quantitative real-time PCR and sequencing surveys demonstrated that the bacterial loads and bacterial profiles of placental tissues were indistinguishable from those of technical controls, thereby complementing the data obtained from the culture component of the study. Additionally, secondary qPCR and 16S rRNA gene sequencing analyses further corroborated the results from the primary analyses briefly described above and further addressed potential concerns over cross-contamination of bacteria from placental tissues to technical controls, and biases introduced through PCR methodology and primer design. Lastly, **as the first**

aim of this chapter, metagenomics was used to evaluate the existence of a placental microbiota in normal term pregnancy. Specifically, metagenomic sequencing was performed on a subset of samples to avoid any potential biases inherent in molecular methods that target specific genes (e.g. 16S rRNA gene) for amplification and/or sequencing. In metagenomics, all the DNA in a sample is sequenced with the absence or only minimal use of PCR. The results from the metagenomic survey reported in this chapter illustrate the importance of using multiple modes of microbiologic inquiry in reevaluating paradigms of sterility: the bacterial signals detected in the placenta were largely inconsistent with those obtained from the 16S rRNA gene surveys, and the bacteria identified were generally ecologically implausible because the placenta would not be a suitable environment in which for them to grow and reproduce.

The second aim of this chapter is to determine if the manner in which 16S rRNA gene sequence data are processed influences the conclusions drawn from sequence-based studies evaluating the existence of a placental microbiota. Since the publication of Theis et al.[150], a review paper was published that included a reanalysis of the data in Theis et al., and in another placental microbiota study by Leiby et al. [167], that used a different sequence processing and classification method [103] than Theis et al. and Leiby et al. had used. Much of the early research on the microbiome using 16S rRNA gene sequencing was performed using operational taxonomic units (OTUs) as the predominant classification method for 16S rRNA gene sequences. Whether OTUs were generated and analyzed using the software programs Qiime2 [190] or mothur [41], the methodology relied on a user-defined sequence similarity identity threshold (97% was standard, although other thresholds such as 95% or 99% were occasionally used), which would group or cluster quality-filtered sequences with $\geq$ 97% sequence identity into a single OTU corresponding to a representative 16S rRNA gene sequence

that is then assigned a specific taxonomic identity/classification. This sequence clustering approach based on sequence identity can account for sequencing errors that are inevitable with existing sequencing technologies. However, it has been suggested that the resolution provided by this clustering method, while valuable for addressing sequencing errors, may obscure microbial biodiversity patterns in biological samples by disregarding potential species and strain level differences that can result from as little as a single nucleotide difference in the 16S rRNA gene [152, 191, 192]. While this issue can be partially ameliorated through OTU clustering at more stringent sequence identity thresholds, such as 99%, which only allows 1 or 2 nucleotide base differences across a 250 base amplicon, there are two additional concerns with the OTU classification method. First, with OTUs, the taxonomic units are constructed based upon the entire dataset, which can require substantial computational time and resources. Second, OTUs are typically generated *de novo*; OTUs from a dataset are specific to that dataset and can therefore not be directly compared across studies or generated based on a reference database, which can limit the taxa that are classified based on the thoroughness of the database [192].

While the majority of microbiome studies have relied on the OTU classification method, an alternative method was recently proposed that identifies exact sequence variants, or amplicon sequence variants (ASVs), thereby providing resolution in the classification of sequences that differ by only a single nucleotide. This classification method, which is performed through the DADA2 [152] package in R [193], addresses the aforementioned caveats associated with OTU approaches and may allow for elucidating underlying patterns in microbial biodiversity that would otherwise be overlooked by OTU classification methods. The ASV approach has been proposed as a more accurate and biologically informative way to classify 16S rRNA gene sequences, specifically in regard to sensitivity and precision [97]. The strengths of this sequence

processing pipeline, in addition to the resolution of sequence variants at a single nucleotide difference, is that ASVs are inferred per sample, rather than per dataset, and are generated based on their relative abundance within a sample. By factoring in relative abundance, the ASV approach excludes extremely low abundance sequences based on the assumption that legitimate sequences are going to occur more frequently than sequences that were resultant from sequencing error.

Proponents of the existence of a placental microbiota have capitalized on the ASV-approach and argued that prior studies that used OTU-based approaches and determined that the bacterial profiles of placental tissues and background technical controls were indistinguishable lacked the resolution required to determine if the molecular signals of bacteria from tissues and controls were indeed from the same microorganisms [103, 175]. In the review by O'Callaghan [103], which re-classified publicly available 16S rRNA gene sequence datasets from the two aforementioned placental microbiota studies [150, 167] using the ASV pipeline, it was suggested that the results and conclusions of the re-analysis were inconsistent with what had been previously published using OTU classifications. Therefore, in this chapter, alongside the analysis of the metagenomic data of placental tissue and technical control samples, I will analyze, compare, and discuss the 16S rRNA gene sequence data from these samples processed using both sequence classification methods. Specifically, I will re-assess the publicly available dataset associated with Theis et al. [150], evaluating the data using ASV and OTU classification methods, compare the data from both methods to illustrate the differences and similarities between them in the context of the original publication, and address several points raised by O'Callaghan et al. [103] that are inconsistent with our analyses (such as their reporting of a high relative abundance of *Ureaplasma* sequences in our negative controls). While the published

manuscript [150] included several datasets generated using multiple PCR methodologies, the reanalysis presented in this chapter is focused on the primary nested PCR dataset and the secondary touchdown PCR dataset, as these were the two amplification methods that yielded large numbers of quality 16S rRNA gene sequences and were not entirely dominated by *Escherichia* sequences, a well-established background DNA contaminant in sequence-based studies [99, 119, 150]. The re-analysis of the data presented here counters the points raised by O'Callaghan et al. [103] and validates our initial findings and conclusions that there is a lack of evidence of a placental microbiota in normal term pregnancies.

III. Methods and Materials

Methods and Materials are further detailed in the published manuscript "Does the human placenta delivered at term have a microbiota? Results of cultivation, quantitative real-time PCR, 16S rRNA gene sequencing, and metagenomics." [150].

Briefly, tissue samples were taken from the placentas of 29 women delivering by cesarean section without labor at term. Samples were collected and processed with aseptic techniques to avoid contamination. Following delivery, placentas were processed in a biological safety cabinet, wherein a core of tissue (i.e. from the amnion through to the basal plate) was collected halfway between the edge of the placental disc and the umbilical cord insertion site. The tissues were placed into sterile containers and frozen at -80° C until DNA extractions were performed.

DNA extraction from placental tissues

During the DNA extraction process, study personnel wore sterile surgical gowns, gloves, and masks, and used individually packaged, sterile, and disposable scalpels and forceps. For each

placental tissue specimen, the amnion and chorion (AC) were collectively cut away from the villous tree and basal plate (V); no tissue components were excluded and both tissue fractions likely contained subchorionic tissue. Genomic DNA was extracted from these two tissue fractions separately. Specifically, DNA was extracted from placental tissues (0.1 to 0.2 g) and background technical control samples using the MoBio PowerSoil DNA Isolation Kit (MoBio Laboratories, 12888), according to the manufacturer's protocol. The DNA extraction kit used, and the mass of placental tissue from which DNA was extracted, were consistent with prior studies of potential placental microbiota [44, 128]. Background technical control samples included extractions performed on: 1) DNA extraction kits without any introduced placental tissue, processed exactly as the placental samples (N = 6); 2) extraction kits whose bead tubes had been exposed to a biological safety cabinet for 20 minutes during placental biopsy collection or processing (N = 16 samples from 3 biosafety cabinets); and 3) extraction kits whose bead tubes had been exposed for 20 minutes to an operating room or microbiology laboratory utilized in this study (N = 21 samples from three operating rooms and three laboratories). These control samples therefore represented five or six technical controls reflecting each potential source of background DNA contamination (i.e. extraction kits, 3 biosafety cabinets, 3 laboratories, and operating rooms), with the three contiguous operating room environments being treated as a single potential contamination source. DNA concentrations of placental tissue and background technical control samples were 42.0 ± 18.5 (standard deviation) ng/µl and ≤ 0.03 ng/µl, respectively. Purified DNA was stored at −20° C.

To eliminate the possibility of any bacterial signal coming from the controls being due to cross-contamination from microbial-populated placental tissues, a secondary series of extractions and analyses were performed that included background technical control samples processed

alongside placental samples and independently of placental samples. Secondary DNA extractions were performed on the collective villous tree & basal plate portion of each of the 29 placental samples. The extraction protocol was the same as that described earlier, except that at least 4 background technical controls were included in each of four rounds of extractions of the placental samples. Specifically, in the first three rounds of extractions, we processed eight placental and four technical control samples. In the fourth round, we processed five placental and five technical control samples. Additionally, we completed a fifth round of extractions composed entirely of 12 blank extraction kit controls, which were not exposed to the atmospheres of the biologic safety cabinets or the laboratories; they were processed exactly as the placental samples. DNA concentrations of placental tissue and background technical control samples were 56.0 ± 24.3 ng/μl and ≤ 0.03 ng/μl, respectively. Purified DNA was stored at −20° C.

Metagenomic sequencing of extracted DNA from placental tissue and background technical control samples

In contrast to sequencing surveys targeting a specific bacterial gene (e.g. 16S rRNA gene), a metagenomic survey entails sequencing all the genes in a clinical sample and assigning the protein-coding genes of bacterial origin to particular bacterial taxa, even potentially at the species level. Nine placental and 11 technical control samples were submitted for metagenomic sequencing using the Illumina HiSeq 4000, 150 paired-end protocol at the University of Michigan's DNA Sequencing Core (Ann Arbor, MI). The placental samples included amnion-chorion and villous tissue and basal plate samples from each of four subjects (subjects 14, 15, 22, and 30), and a villous tissue and basal plate placental sample from one subject (subject 19). The technical control samples included eight biological safety cabinet and three blank extraction kit samples.

Metagenomic sequence data processing using MG-RAST

The twenty metagenomic sequence libraries were submitted to the MG-RAST metagenomes analysis server at the Argonne National Laboratory [194]. Forward and reverse reads were combined into joined paired-end reads, as applicable, yet those with non-overlapping paired-ends were retained as well. Default pipeline parameter options were used: assembled (no), dereplication (yes), and screening (*H. sapiens*, NCBI v36). Dynamic trimming was also set to default, except lowest quality base phred score was set to 10, and sequences were trimmed to contain at most seven bases with a phred score of 10 [195]. Reads more than two standard deviations from the mean read length were discarded, as were poor quality and artificial duplicate reads [194]. As the final processing step, sequences were screened for host DNA, in this case those that matched *H. sapiens*, via NCBI v36. SortMeRNA [196] was used to identify rRNA genes with a 70% identity cut-off and CD-HIT [197] was used to cluster those with a 97% nucleotide similarity. The longest representative from these clusters was run through a BLAST-like alignment tool (BLAT) [198] similarity search against the M5rna database for rRNA identification [194]. Sequences of potential protein coding regions were identified via FragGeneScan [199], and predicted protein coding sequences were clustered at 90% identity with CD-HIT and run through a BLAT search against the M5NR protein database. Protein features were excluded if they overlapped with ribosomal RNA features. Identified rRNA and protein sequences were annotated and mapped back to the original sequences. Taxonomic assignments were made using the GenBank database and the default MG-RAST parameters: maximum e-value cutoff of 5, minimum percent identity cutoff of 60%, minimum alignment length cutoff of 15, minimum abundance of 1, and representative-hit classification [194]. For gene function characterization, sequences were mapped to the KEGG Orthology (KO) database

[200]. Metagenomic sequencing files are publicly available at MG-RAST (Project ID PRJNA397876).

Reanalysis of published placental data after reclassification into ASVs

ASV sequence data processing using DADA2

Classification of ASVs was performed as described by Callahan et al. [152] using the DADA2 tutorial pipeline (benjjneb.github.io/dada2/tutorial.html). The tutorial was followed using the default parameters except in the following instances: multithreading was changed to FALSE throughout the pipeline; for the filter and trimming step the truncation length for forward reads was increased from 240 to 250 ("truncLen=c(**250**,160)"), and the maximum number of expected errors was increased from 2 to 7 for reverse reads ("maxEE=c(2,**7**)"); for the removal of chimeras, the method was changed to "**pooled**"; and to assign taxonomy, a minimum bootstrapping value was added ("**minBoot=80**") and the reference database for assigning taxonomy was changed from the default Silva reference database to the RefSeq RDP 16S database v2 May 2018, as indicated in the review by O'Callaghan et. al. [103] and the associated GitHub repository (https://github.com/jessicaocallaghan/reproductive_microbiome/blob/master/assign_taxonomy_script).

Following DADA2 processing, ASVs identified as chloroplasts were removed from the primary nested (n = 2 ASVs) and secondary touchdown (n = 6 ASVs) datasets. After removing the chloroplast ASVs, the primary nested dataset contained 5,239,414 reads and 714 ASVs, and the secondary touchdown dataset contained 2,050,376 reads and 704 ASVs. A preliminary analysis identified a significant proportion of human mitochondrial sequences in the secondary

touchdown dataset; after these sequences were removed, there were 1,642,769 sequences and 492 ASVs in the secondary touchdown dataset.

Subsampling of 16S rRNA gene sequence data for ASV vs OTU comparisons

Both primary nested and secondary touchdown datasets were subsampled to a sequence depth of 500, unless otherwise indicated, to mirror the analyses performed by O'Callaghan et al. [103]. In order to ensure that the bacterial profiles of samples remained sufficiently characterized after subsampling, samples were only included in analyses if they had a Good's coverage ≥ 95%.

Decontam program for the removal of likely background DNA contaminants

Decontam is a program that classifies OTUs or ASVs in biological samples as either contaminants or non-contaminants based upon their distribution among complementary blank DNA extraction controls [120]. The decontam package was run in R Studio [201] and our ASV-specific analyses were done using two methods. The *isContaminant* method was used for both primary nested and secondary touchdown datasets and parameters were kept in accordance with those used in O'Callaghan et al. [66]: the prevalence method was used with a threshold of 0.5 (default is 0.1). The alternative *isNotContaminant* method was also used for both datasets, using the prevalence method and thresholds of 0.5 and 0.6 for the primary nested dataset and 0.2 and 0.5 for the secondary touchdown dataset (default is 0.5). For each analysis, all background technical control sample types were grouped together as controls in the metadata file.

Rationale for use of isNotContaminant method and thresholds

The alternative method, *isNotContaminant*, in the *decontam* package is the most appropriate method for using in studies with low microbial biomass samples [120]. This method begins with the assumption that all sequences are contaminants and that the presence of an ASV

in a greater proportion of biological samples than controls identifies the sequence as a non-contaminant. This method is valuable as it provides greater confidence in describing bacterial communities from sequence data when a large proportion of contaminant sequences are expected. It is not appropriate, however, to assume that taxa that remain after filtering are evidence of a legitimate bacterial community. It is critical that researchers using this tool further investigate whether the remaining taxa are ecologically plausible and are exceedingly and consistently present in biological samples more so than in controls.

Additionally, when using *decontam*, as addressed in Davis et al. [120], it is critical to assess the distribution of decontam scores for each dataset in order to establish a cutoff or "threshold" of when sequences can be considered contaminants. This is important because a particular threshold for contaminant or non-contaminant classification may not be appropriate for every dataset and, in the case of the *isNotContaminant* function, the threshold classifies scores below the designated threshold as non-contaminants (e.g. a threshold of 1.0 would identify all sequences < 1.0 as non-contaminants). To identify appropriate thresholds for the primary nested and secondary touchdown datasets, a histogram was generated for each dataset illustrating the distribution of decontam scores, their presence in placental samples, and the number of ASVs at each score **(Figure 2.1)**. The default threshold is 0.5 for this method and both datasets were analyzed at this threshold. Additionally, while the histogram for the primary nested dataset illustrated that there is a widespread distribution of scores, the large peak at 0.5 can be viewed as an inflection point for this dataset **(Figure 2.1A)**. Because a large number of ASVs had a score between 0.5 and 0.6, we performed an additional analysis through *decontam* with a threshold of 0.6, which would represent a less conservative assessment of the data. The histogram for the

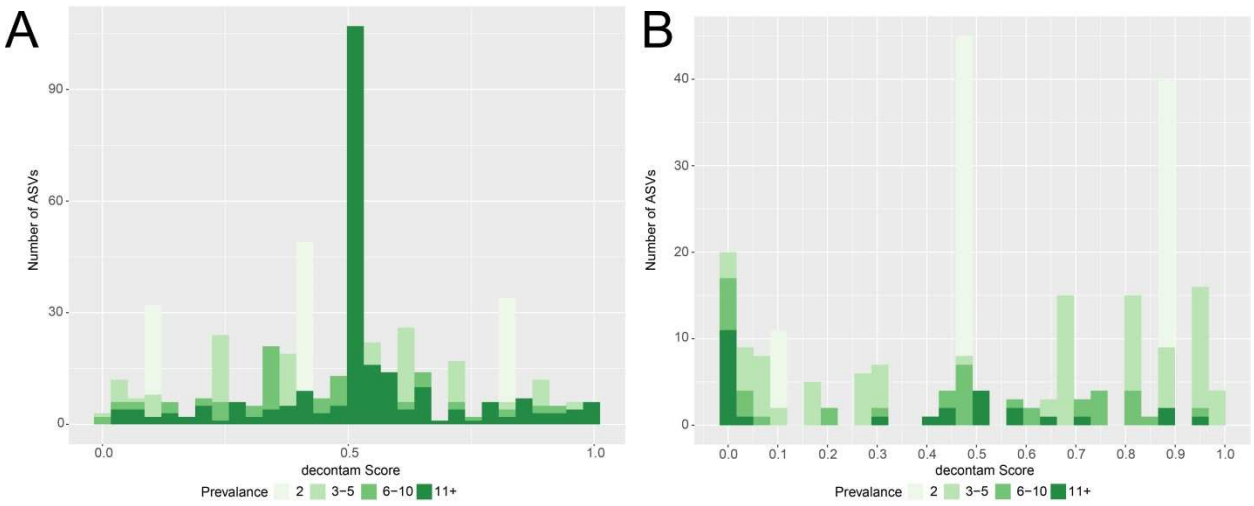

Figure 2.1. Histograms of decontam scores for identifying appropriate thresholds for filtering of the ASV datasets. Decontam scores indicating that A) thresholds of 0.5/0.6 for the primary nested PCR dataset should be appropriate for identifying non-contaminant sequences, and B) thresholds of 0.2/0.5 for the secondary touchdown PCR dataset appear appropriate owing to a more disordered distribution of scores.

secondary touchdown dataset showed several peaks with a much less clear distribution, however, an early cluster of low scores suggested a very conservative threshold of 0.2 could be appropriate for this dataset, in addition to the default 0.5 threshold **(Figure 2.1B)**.

Filtering of contaminant ASVs from the primary nested dataset using decontam

Decontam analysis using isContaminant and a threshold of 0.5

After running *decontam* on the primary nested dataset, 241 of 714 ASVs were identified as "TRUE" contaminants, leaving 66.2% (473/714) of the ASVs in the analysis. After removing the contaminant ASVs, the total sequence reads were reduced from 5,239,414 to 3,792,300 (72.4% of the original dataset). Good's coverage values remained > 98% after contaminant ASVs were removed, however, one placental sample (23V) and three technical controls (3Room1, 3Room4, 4Hood4) contained less than 500 sequences and were removed prior to subsampling. After subsampling, 283 ASVs remained and one placental sample (4V) and three controls (3Hood5, 4Hood2, 7Hood2) with Good's coverage values < 95% were removed from analysis.

Decontam analysis using isNotContaminant and a threshold of 0.5

Applying the 0.5 threshold for the *isNotContaminant* function, the primary nested dataset contained 193 ASVs that were classified as non-contaminants. This represented 27.0% (193/714) of the total ASVs prior to running decontam. After removing the contaminant ASVs, coverage remained above 95% for all samples, however 16 samples no longer had ≥ 500 sequences. Of these 16 samples, two were placental samples (13AC, 23V), three were kit controls, four were hood controls, and seven were room controls (one of which was an operating room control). These samples were removed from analysis and the remaining samples were subsampled to 500

sequences. After subsampling, coverage for all samples remained above 95% and 156 ASVs were represented.

Decontam analysis using isNotContaminant and a threshold of 0.6

Using a less conservative threshold of 0.6, 348 ASVs were classified as non-contaminants. This represented 48.1% (348/714) of the total ASVs prior to running decontam. After removing the contaminant ASVs, Good's coverage values remained > 95% for all samples; however, three samples no longer had ≥ 500 sequences. Of these three samples, two were room controls and one was an extraction kit control. These samples were removed from analysis and the remaining samples were subsampled to 500 sequences. After subsampling, the Good's coverage values for six samples were < 95% (3V, 6V, 18AC, and 3 Hood controls) and these samples were removed from analysis. The analysis proceeded with 235 ASVs represented in the dataset.

Filtering of contaminant ASVs from the secondary touchdown dataset using decontam

To maximize power in discriminating contaminating from non-contaminating ASVs, the secondary touchdown dataset was processed through *decontam* using all kit controls (n = 29, including DNA extraction kits processed alongside as well as independent of placental samples). There was no difference in the bacterial profiles of controls processed alongside placentas and those processed alone (NPMANOVA: $F = 1.069$, $p = 0.278$). Yet, to be conservative, following the execution of *decontam*, the kit controls not processed alongside placental samples were removed from analyses comparing and contrasting the bacterial profiles of control and biological samples. The removed kit controls accounted for 549,488 of the 1,642,769 total sequence reads, and 75 of the 492 total unique ASVs. Placental samples and the remaining controls were retained if they had ≥ 500 reads after the contaminant ASVs were removed (25/29 placental samples and

16/17 control samples). The total read count after removing samples with < 500 reads was 1,091,208 and Good's coverage values were ≥ 99% for all samples.

Decontam analysis using isContaminant and a threshold of 0.5

Of the 417 ASVs left in the dataset, 103 were identified as contaminants leaving 75.3% (314/417) of the ASVs in the analysis. The read count was reduced to 59.3% (647,016/1,091,208) and one control sample (TD_B3_3) was removed for having < 500 reads. All remaining samples had Good's coverage values ≥ 99% and were subsampled to 500 sequence reads. After subsampling, Good's coverage values were> 98% and 277 ASVs remained.

Decontam analysis using isNotContaminant and a threshold of 0.5

For this analysis, 298 ASVs were identified as contaminants, leaving 28.5% (119/417) of the ASVs in the analysis. The read count was reduced to 45.1% (492,014/1,091,208). Only one sample was removed due to having < 500 reads (the same control as above: TD_B3_3). All remaining samples had Good's coverage values ≥ 99% and were subsampled to 500 sequence reads. After subsampling, Good's coverage values were > 98% and 103 ASVs remained.

Statistical analyses

16S rRNA gene profile structure comparisons of the primary nested PCR dataset

Bray-Curtis similarity indices were calculated and differences in bacterial profiles between placental tissue samples and blank DNA extraction controls were statistically evaluated using one-way non-parametric multivariate ANOVA (NPMANOVA) with Bonferroni corrections applied. These analyses were completed using the software program PAST (v. 3.25) [202].

16S rRNA gene profile structure comparisons of the secondary touchdown PCR dataset

Bray-Curtis similarity indices were calculated and differences in bacterial profiles between placental tissue samples and blank DNA extraction controls were statistically evaluated using NPMANOVA in the vegan [203] package in R (v. 3.61) [193] allowing for the analysis of multiple group variables and their interaction.

Figure generation

All Principal Coordinates Analysis (PCoA) plots were generated in PAST (v. 3.25) [202]. All heatmaps were generated using Morpheus [204] software. Prior to heatmap generation, unweighted pair group method with arithmetic mean clustering was performed based on Bray-Curtis similarity values in PAST [202].

IV. Results

The results from metagenomic sequencing presented below were published in the manuscript "Does the human placenta delivered at term have a microbiota? Results of cultivation, quantitative real-time PCR, 16S rRNA gene sequencing, and metagenomics" [150].

Metagenomic surveys of the bacterial profiles of placental tissues

At least 43,000,000 sequence reads were obtained from each of nine placental tissue samples (61,027,678 ± 3,190,738 SEM; **Table 2.1**). On average, 0.05% of these sequences were classified as bacterial in origin. Good's coverage values (99.61% ± 0.14 SEM) indicated that the bacterial profiles of these samples were thoroughly characterized from a taxonomic standpoint. The survey identified 267 bacterial genera, with 19 being considered prominent, defined as having an average relative abundance of ≥ 0.1% (**Figure 2.2)**. Only five genera had an average relative abundance ≥ 1.0%: *Cyanothece*, *Coprobacillus*, Candidatus Phytoplasma, *Chlorobium*,

Table 2.1. Summary of metagenomic sequence data from placental samples.

Placenta	Total sequences	% removed after quality control	GenBank hits	Bacterial hits	% of hits that were bacterial	Bacterial genera per sample	Good's coverage	# of these genera detected in controls		
								All	5	< 5
14AC	77,603,776	95.63	350,330	40,188	11.47	126	99.8%	125	1	0
14V	63,783,888	95.49	291,868	33,300	11.41	116	99.8%	116	0	0
15AC	60,438,600	95.58	287,700	32,400	11.26	106	99.8%	106	0	0
15V	62,928,591	95.63	278,408	34,638	12.44	112	99.8%	112	0	0
19V	62,542,939	95.62	232,037	34,980	15.08	99	99.8%	99	0	0
22AC	68,611,076	95.37	333,864	37,358	11.19	116	99.8%	116	0	0
22V	52,974,105	95.65	236,102	28,788	12.19	97	99.8%	97	0	0
30AC	56,771,638	95.43	266,375	8,353	3.14	124	98.5%	124	0	0
30V	43,594,487	95.11	249,226	11,309	4.54	92	99.5%	92	0	0
Average	**61,027,678**	**96**	**280,657**	**29,035**	**10.30**	**110**	**99.6%**			

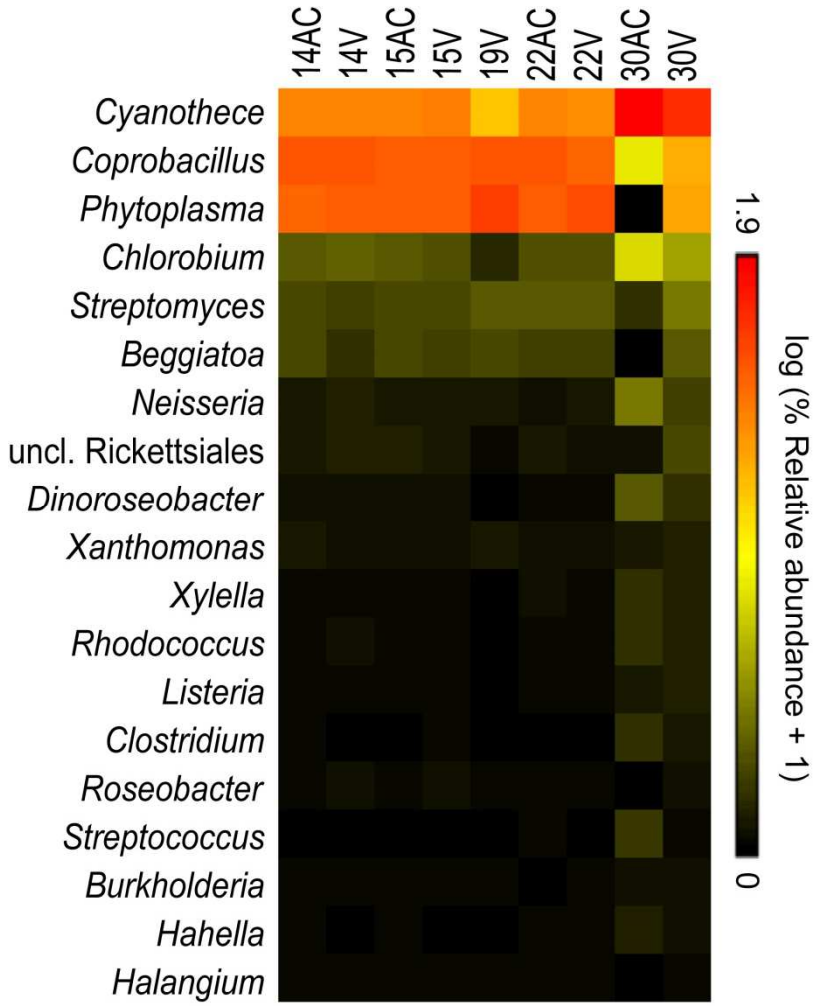

Figure 2.2. Bacterial profiles of placental samples as determined by metagenomic sequencing. Heatmap illustrating the relative abundances of prominent bacterial genera among placental samples. Prominent genera were here defined as those having an average relative abundance ≥ 0.1% among placental samples. AC indicates amnion and chorionic plate samples, and V indicates samples of the villous tree and basal plate.

and *Streptomyces*. *Escherichia* was present in each placental sample, with an average relative abundance of 0.05%. A small fraction (0.1%) of the bacterial reads were confidently assigned a gene function **(Table 2.2)**. Broadly, these bacterial gene functions were metabolism (amino acid, carbohydrate, vitamin and energy metabolism), genetic machinery (DNA translation, replication, repair and degradation), and environmental processing (membrane transport and signal transduction), all of which were also heavily represented in the bacterial gene function profiles of the technical controls **(Table 2.3)**.

Given the necessary differences in metagenomic library preparation for the placental tissue and technical control samples, their broad bacterial profiles cannot be compared in a quantitative manner. However, it is reasonable to inquire if there are genera consistently identified in placental tissue samples that were not also widely present in the sequenced background technical controls. There were 36 genera present in all nine sequenced placental tissue samples, and 89 genera present in at least half of these samples. Each of these 125 genera was present in all 11 sequenced background technical controls. Of the 267 total genera, or approximate genus-level taxa, identified in placental tissue samples, only one was not found in every control sample **(Table 2.1)**: an unclassified Myxococcales, present in one placental sample with a relative abundance < 0.01%.

Of the prominent genera (≥ 1% average relative abundance) identified in the primary 16S rRNA gene sequencing analysis of this study **(Figure 2.3)**, *Clostridium* and *Propionibacterium* were detected in each of the nine placental samples via metagenomic sampling. *Staphylococcus* was present in 8/9, *Stenotrophomonas* was present in 6/9, *Achromobacter* was present in 5/9, *Methylobacterium* and *Paracoccus* were present in 3/9, *Acinetobacter* was present in 2/9, and

Table 2.2. Summary of metagenomic sequence reads from placentas assigned a bacterial gene function.

KEGG Orthology	14AC	14V	15AC	15V	19V	22AC	22V	30AC	30V
Amino Acid Metabolism	11	12	10	11	10	2	7	28	4
Biosynthesis of Other Secondary Metabolites	0	0	0	0	0	0	0	2	0
Carbohydrate Metabolism	4	8	3	3	0	2	0	3	2
Cell growth and death	0		0	2	0	1	0	0	0
Cell motility	0	0	3	0	0	1	4	1	0
Energy Metabolism	0	2	0	2	2	0	1	10	0
Folding, sorting and degradation	4	1	0	2	0	0	0	2	1
Glycan biosynthesis and metabolism	1	0	0	0	0	0	0	1	1
Lipid metabolism	0	0	0	5	2	0	0	0	0
Membrane Transport	4	11	8	1	7	9	3	7	3
Metabolism of Cofactors and Vitamins	6	2	0	2	5	0	1	1	2
Metabolism of Other Amino Acids	1	0	0	0	0	0	0	0	0
Metabolism of Terpenoids and Polyketides	0	0	0	1	0	0	0	0	0
Nucleotide metabolism	0	1	0	2	0	0	0	3	0
Replication and repair	0	0	1	0	1	2	2	15	4
Signal transduction	2	2	2	1	0	3	0	3	1
Transcription	0	0	0	0	0	0	0	2	0
Translation	3	3	1	0	0	1	1	8	1
Transport and catabolism	2	0	0	0	0	3	0	2	0
Xenobiotics Biodegradation and Metabolism	0	2	0	0	2	0	0	0	0

Table 2.3. Summary of metagenomic sequence data from technical controls.

Control	Total sequences	% removed after quality control	GenBank hits	Bacterial hits	% of hits that were bacterial	Bacterial genera per sample	Good's coverage	Average # of bacterial reads assigned to a gene function
4hood1	47,235,826	90.26	604,070	425,500	70.44	604	100.0%	128,294
4hood2	55,527,650	92.52	1,291,131	1,189,969	92.16	630	100.0%	428,646
4hood3	53,435,968	94.75	980,709	871,958	88.91	626	100.0%	308,573
4hood4	55,516,701	95.06	988,718	885,749	89.59	623	100.0%	295,312
7hood1	60,045,409	95.49	349,714	207,703	59.39	608	99.9%	70,931
7hood2	71,332,222	95.09	622,774	463,449	74.42	636	99.9%	161,243
7hood3	58,712,747	94.65	336,649	174,036	51.70	594	99.9%	56,894
7hood5	62,886,028	94.52	1,069,227	916,346	85.70	636	100.0%	323,429
Kit4	36,260,902	90.21	441,464	345,654	78.30	623	99.9%	118,668
Kit5	38,429,367	93.72	744,139	657,313	88.33	629	100.0%	230,137
Kit6	45,929,752	94.24	635,400	528,815	83.23	636	100.0%	192,513
Average	**53,210,234**	**93.68**	**733,090**	**606,045**	**78.38**	**1750**	**100.0%**	**210,422**

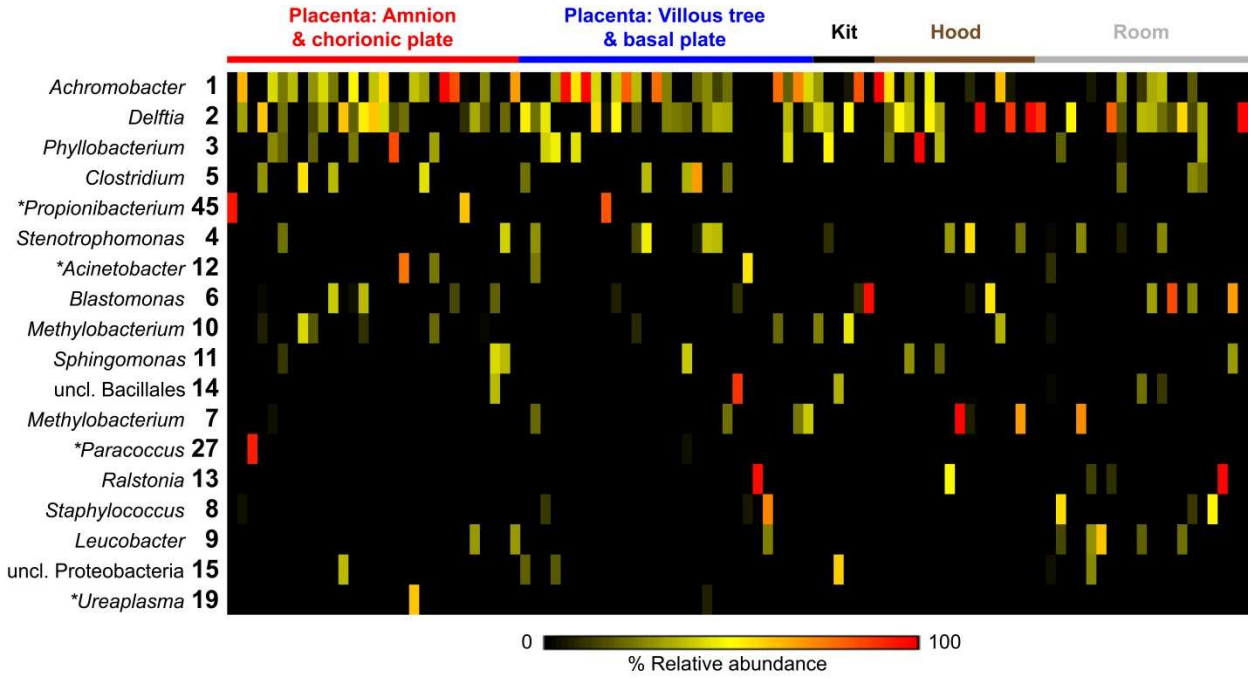

Figure 2.3. Bacterial profiles of placental samples based on 16S rRNA gene sequencing data. Heatmap illustrating similarity in percent relative abundances of prominent operational taxonomic units (OTUs) among placental samples and technical controls. Prominent OTUs were defined as those having an average relative abundance ≥ 1% among the placental samples. OTUs were generated using a 97% sequence similarity cutoff and the primary nested PCR data set. Asterisks indicate OTUs that were prominent in placental samples but not in technical controls.

Delftia and *Ureaplasma* were present in 1/9. Of these genera, only *Clostridium* was present in placental metagenomic bacterial profiles at an average relative abundance ≥ 0.1% **(Figure 2.2)**.

Bacterial profiles of placental tissues and technical controls from the primary nested PCR dataset characterized using OTU and ASV approaches

The results for the primary nested PCR analysis did not change after reanalyzing the sequence data using the ASV approach **(Table 2.4, Figure 2.4)**. Both statistically and visually, the bacterial profiles of placental samples and technical controls were indistinguishable using either classification method **(Figure 2.4, Figure 2.5A)**.

Bacterial profiles of placental tissues from the primary nested PCR dataset after removing any ASVs that were identified in controls

The primary nested PCR dataset included 714 ASVs and had average read counts for amnion-chorion and villous tree & basal plate samples of 48,961 and 51,631, respectively **(Figure 2.6)**. After removing any ASVs that were identified in technical controls, the average read counts for amnion-chorion and villous tree & basal plate samples dropped to 189 and 154, respectively **(Figure 2.6)**. Only 118/714 (16.5%) ASVs from the full dataset remained, and these ASVs were represented by only 0.3% of the total reads from placental samples **(Figure 2.7)**. After removing these 596 ASVs identified in controls, 9951 reads remained in the dataset leaving eight amnion-chorion and 10 villous tree & basal plate samples with no reads. After removing placental samples with no reads, the average read count for amnion-chorion samples (n = 21) was 261; the average read count for villous tree & basal plate samples (n = 19) was 236. Of the remaining 118 ASVs, no ASV was detected in more than four placental samples. The only two ASVs detected in four placental samples were classified within the bacterial families Alcaligenaceae and Caulobacteraceae. Notably, the three *Ureaplasma* ASVs in the original dataset were also removed due to occurrence in controls (4, 1, and 1 reads total for the 3

Table 2.4. Comparison of the bacterial profiles of placental samples and technical controls using OTU and ASV approaches for the primary nested PCR dataset. ASV and the original OTU sample datasets were subsampled to 500 sequences each and only samples with ≥ 95% Good's coverage values were included. Non-parametric multivariate analysis of variance (NPMANOVA) based on the Bray-Curtis index was used to assess variation among placental samples and technical controls. Results of overall global effect analyses are presented along with the results of pairwise comparisons that involve placental samples. Probability values for these permutation tests were not adjusted for multiple pairwise comparisons, because this can be overly conservative. However, for pairwise tests that were statistically significant, we present the Bonferroni corrected probability value in parentheses.

		Structure (OTUs)		Structure (ASVs)	
		F	**p-value**	**F**	**p-value**
Amnion-chorion	**Global**	**1.153**	**.244**	**1.126**	**.275**
	Rooms	2.275	.010 (.058)	2.018	.021 (.128)
	Hoods	1.228	.256	0.585	.391
	Kits	0.530	.885	1.353	.907
Villous tree & basal plate	**Global**	**1.217**	**.188**	**1.289**	**.129**
	Rooms	2.489	.008 (.045)	1.463	.005 (.028)
	Hoods	1.076	.359	0.670	.859
	Kits	0.829	.580	1.243	.082

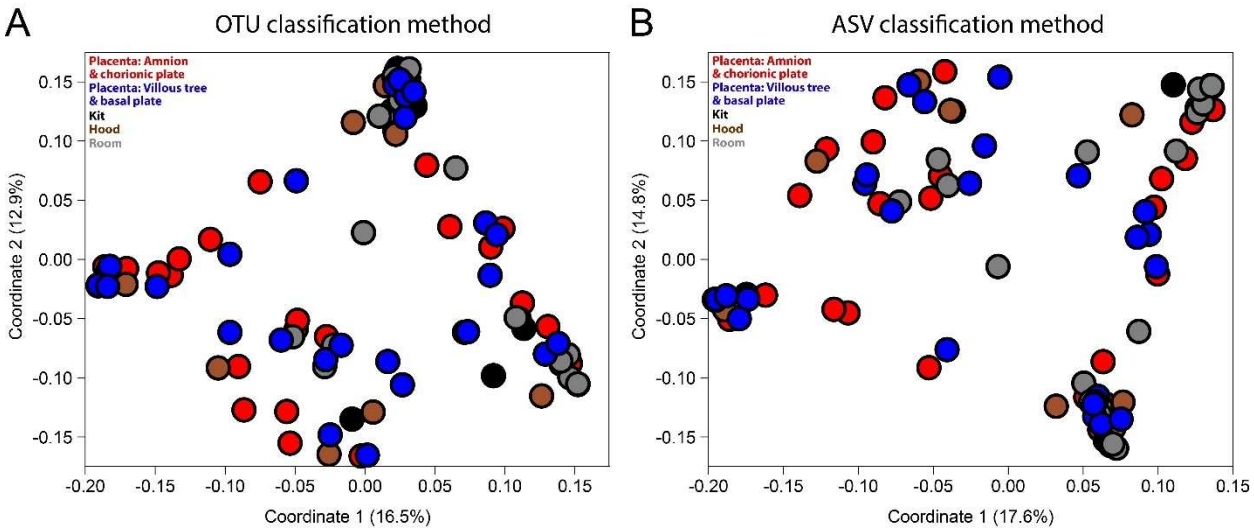

Figure 2.4. Principal Coordinates Analysis (PCoA) comparing the bacterial community structure of OTU and ASV approaches for the primary nested dataset. For both datasets, placental samples and controls overlap illustrating the lack of variation between sample types.

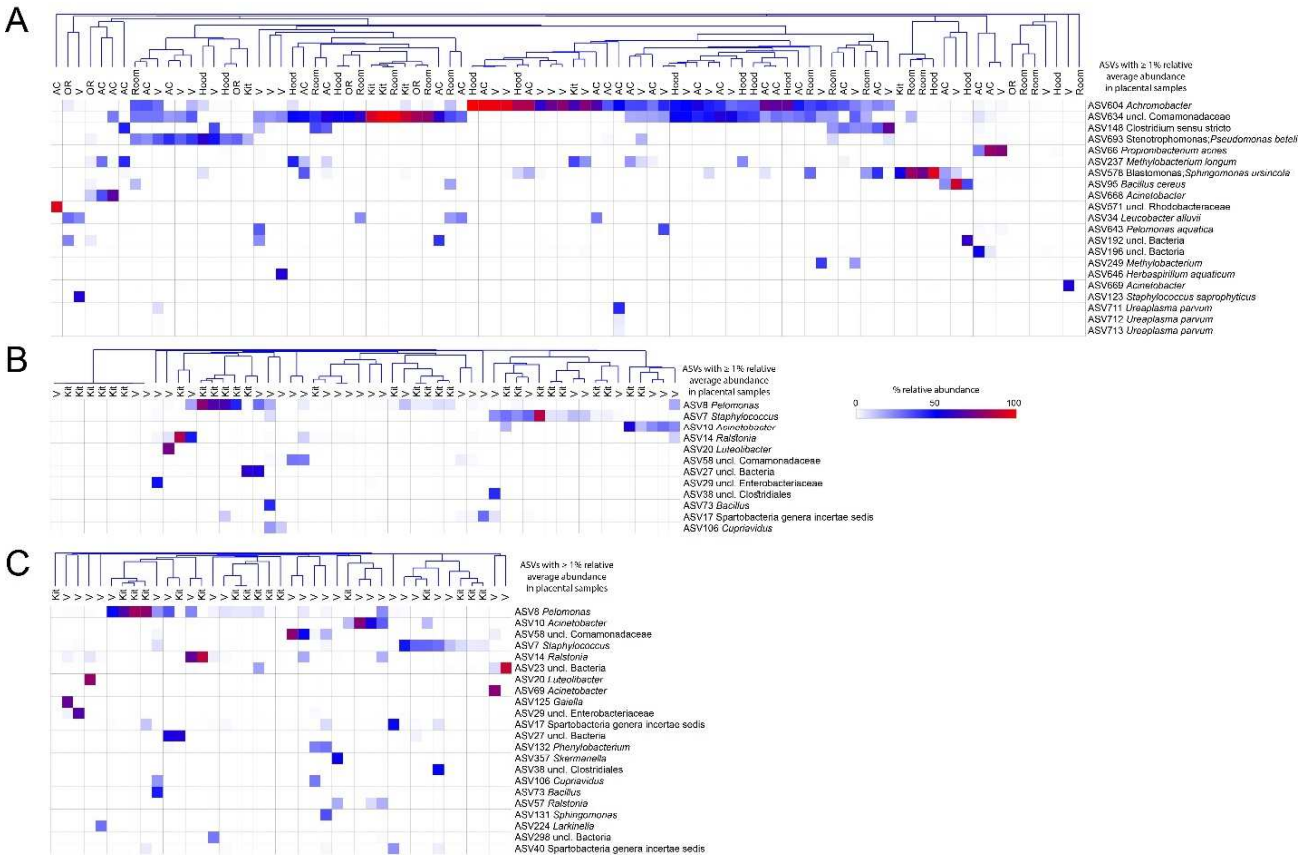

Figure 2.5. Heatmaps of the bacterial profiles from ASV datasets. Heatmaps of placental datasets with ASVs ≥ 1% relative abundance illustrating that placental samples cluster indiscriminately with control samples. A) Primary nested dataset corresponding to the dataset indicated in Figure 2.3B. All Ureaplasma ASVs were included despite not having an average relative abundance ≥ 1% in placental samples. B) Secondary touchdown dataset with no ASV removal and from the same dataset illustrated in Figure 2.10B. Panel C) Secondary touchdown dataset with mitochondrial ASVs removed and from the same dataset illustrated in Figure 2.10C. Samples for all panels were clustered by hierarchical clustering using Bray-Curtis similarity indices based on the ASVs on the Y-axis of each heatmap.

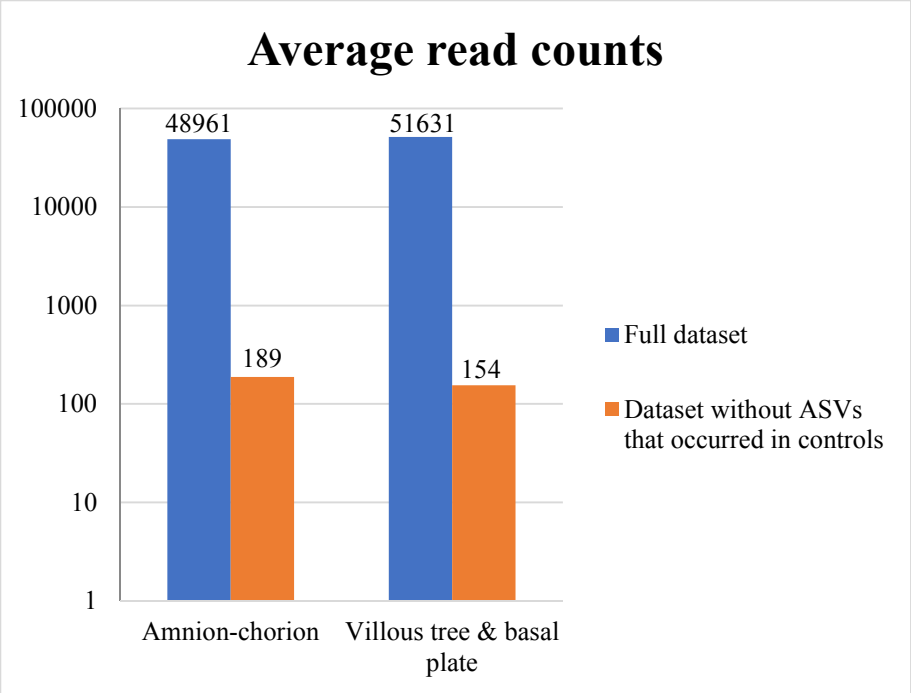

Figure 2.6. Average sequence read count for placental samples from the primary nested PCR dataset. Read counts include before and after removal of ASVs that occurred in controls. The Y-axis is graphed on a log10 scale.

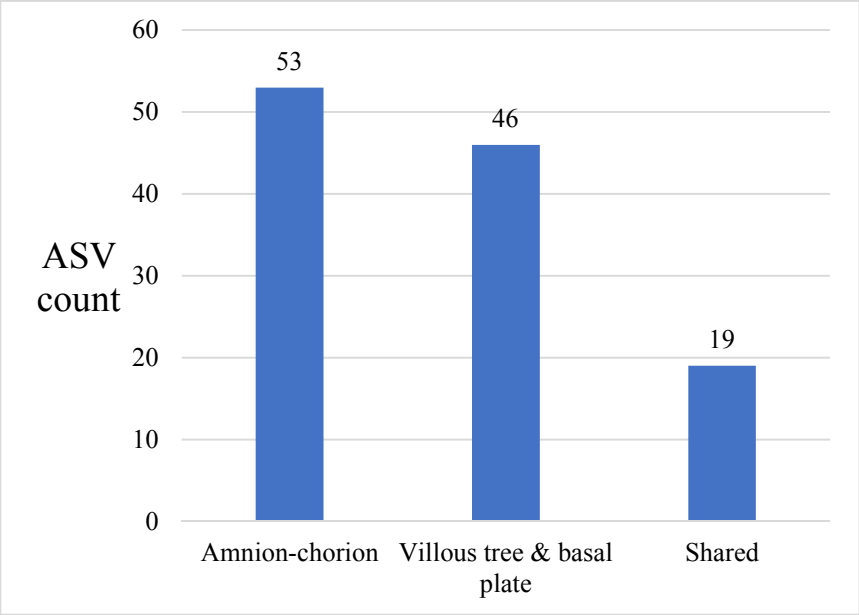

Figure 2.7. Unique ASV counts per placental sample type after removal of ASVs that were identified in control samples.

Ureaplasma ASVs). Only 18 placental samples had > 100 reads and only 12 samples had > 250 reads. Good's coverage at either sequencing depth exceeded 97%, yet a subsequent taxonomic analysis was only performed on placental samples with ≥ 250 remaining reads. Principal Coordinates Analyses (PCoAs) could not be generated for this dataset because only one ASV was detected in more than one sample. This was the same Caulobacteraceae ASV reported above; it occurred in two villous tree & basal plate samples at 942 reads (942/942 total reads) and one read (1/651 total reads), respectively, in these two samples. The limited magnitude of a remaining bacterial signal in placental samples, combined with a lack of consistency in the ASVs detected across samples, suggests that the removal of all ASVs identified in technical controls from the dataset before the characterization of a potential placental microbiota is not an appropriate approach. Removing these data essentially removes any analyzable bacterial signal.

Bacterial profiles of placental tissues from the primary nested PCR dataset after filtering out likely background DNA contaminants identified through the program decontam

After filtering the primary nested PCR dataset through *decontam* and removing the ASVs identified as contaminants, statistical analysis of the structure of the bacterial profiles of placental samples and technical controls did not change compared to the unfiltered dataset **(Table 2.5, Figure 2.8)**, regardless of the *decontam* approach used. In fact, in some cases, p-values of statistical analyses actually increased when compared to the analyses without using decontam, further supporting the conclusion that the bacterial profiles of placental samples and technical controls are not distinct **(Table 2.5, Figure 2.9)**.

Bacterial profiles of placental tissues and technical controls from the secondary touchdown PCR dataset characterized using OTU and ASV approaches

After the reclassification of 16S rRNA gene sequence data into ASVs, the secondary touchdown PCR analysis did initially appear to reveal an effect of sample type, round of

Table 2.5. Comparison of the structure of bacterial profiles of placental samples and technical controls before and after *decontam* filtering for the primary nested PCR dataset. The values on the left are identical to those in the right panel in Table 2.4. The other panels are evaluating variation between placental samples and controls after applying three different filtering methods through *decontam*. In the second panel, the *isContaminant* function in decontam was used. In the third and fourth panels, the conservative *isNotContaminant* function was used at two different thresholds. All sample datasets were subsampled to 500 sequences (samples were excluded if they had < 500 reads) and only samples ≥ 95% Good's coverage values were included. Non-parametric multivariate analysis of variance (NPMANOVA) based on the Bray-Curtis index was used to assess variation among placental samples and technical controls. Results of overall global effect analyses are presented along with the results of pairwise comparisons that involve placental samples. Probability values for these permutation tests were not adjusted for multiple pairwise comparisons, because this can be overly conservative. However, for pairwise tests that were statistically significant, we present the Bonferroni corrected probability value in parentheses.

		Bacterial community structure							
		ASVs		***Decontam* filtered**					
				***isContaminant* 0.5 threshold**		***isNotContaminant* 0.5 threshold**		***isNotContaminant* 0.6 threshold**	
		F	**p-value**	**F**	**p-value**	**F**	**p-value**	**F**	**p-value**
Amnion -chorion	**Global**	**1.126**	**.275**	**1.033**	**.405**	**0.844**	**.655**	**1.318**	**.147**
	Rooms	2.018	.021 (.128)	1.883	.070	0.847	.537	2.558	.012 (.071)
	Hoods	0.585	.391	0.636	.736	0.968	.431	0.621	.790
	Kits	1.353	.907	1.138	.313	0.841	.528	1.416	.193
Villous tissue & basal plate	**Global**	**1.289**	**.129**	**1.227**	**.208**	**0.782**	**.752**	**1.405**	**.087**
	Rooms	1.463	.005 (.028)	2.385	.024 (.142)	.649	.784	2.775	.005 (.029)
	Hoods	0.670	.859	0.785	.593	.853	.512	0.623	.817
	Kits	1.243	.082	1.414	.175	.916	.437	1.617	.110

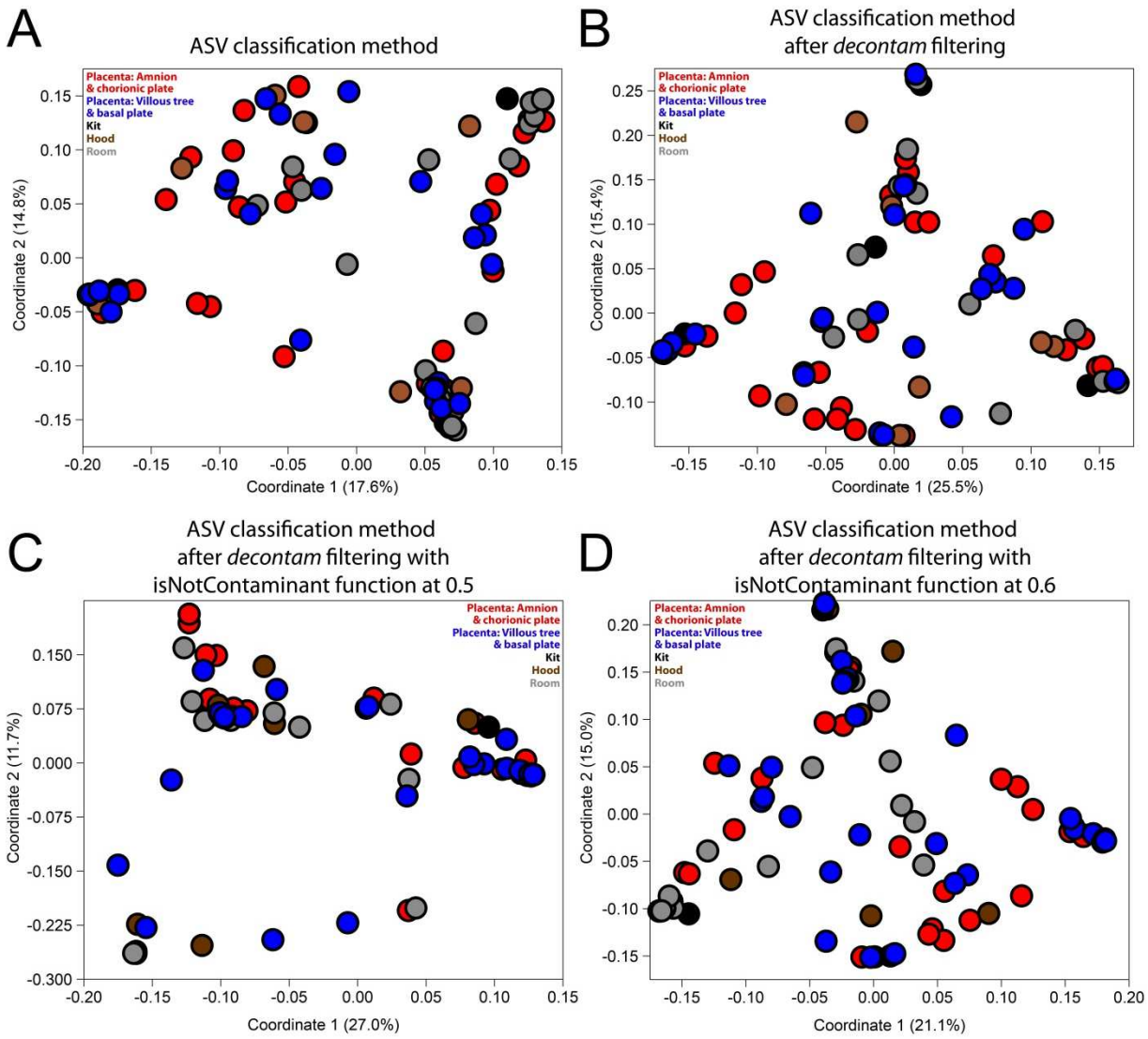

Figure 2.8. Principal Coordinates Analyses (PCoAs) illustrating the structure of the bacterial profiles of placental samples and technical controls from the primary nested PCR dataset after various contaminant filtering methods. Panel A is the same as Figure 2.4B. In panel B, ASVs identified as contaminants by the *isContaminant* function in *decontam* were removed. In panels C and D, two different thresholds were applied to identify and subsequently remove contaminants using the alternative function *isNotContaminant* in the *decontam* package. For all panels, sample datasets were subsampled to 500 sequence reads (samples were excluded if they had < 500 reads) and only samples with ≥ 95% Good's coverage values were included.

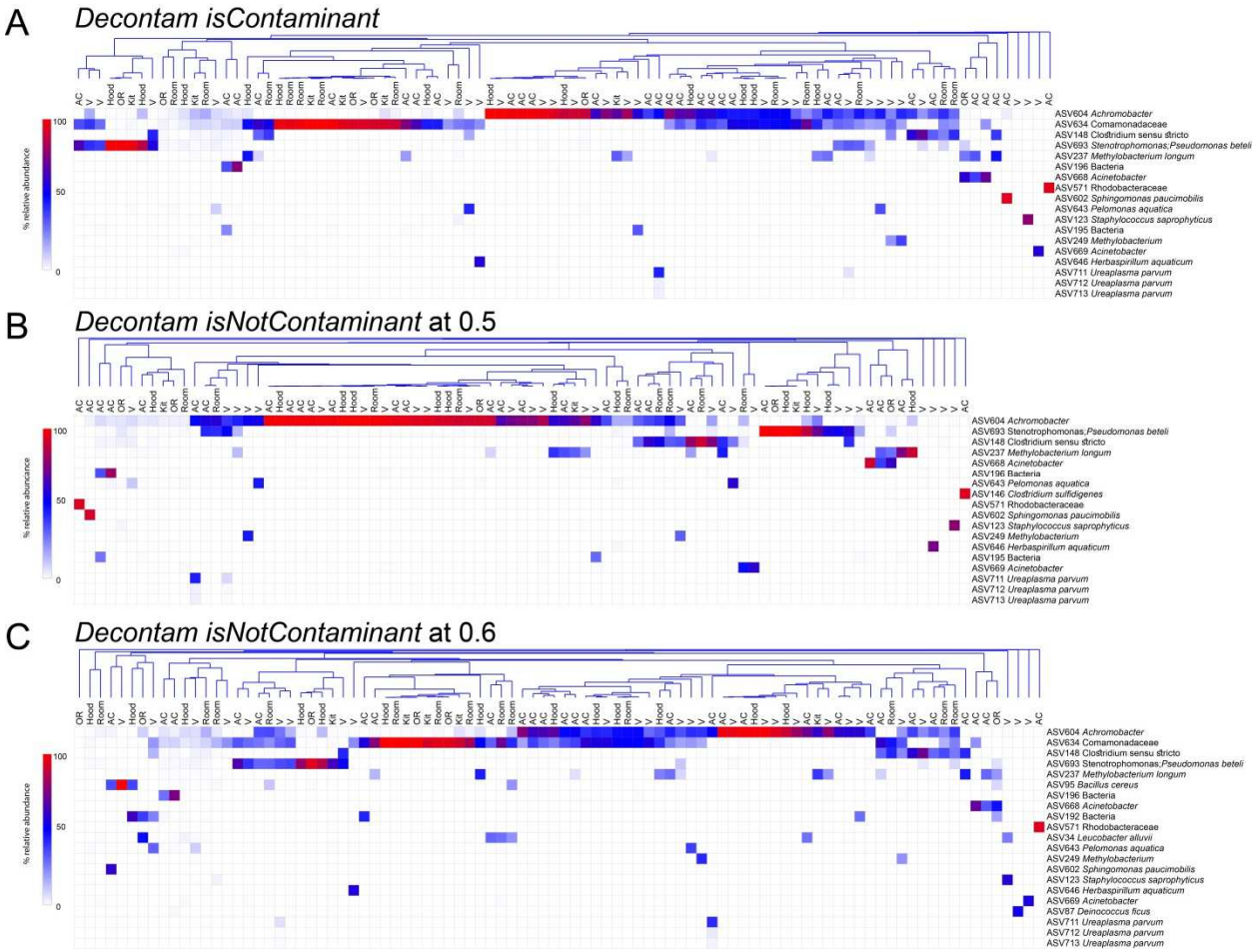

Figure 2.9. Heatmaps of the primary nested PCR dataset after decontam filtering. Bacterial profiles of placental samples and technical controls cluster indiscriminately, regardless of filtering method**.** The Y-axis indicates ASVs with ≥ 1% average relative abundance across placental samples for each respective dataset; also included are *Ureaplasma* ASVs. Samples are clustered on the x-axis by Bray-Curtis similarities.

extraction, and an interaction effect between the two, effects which were not evident in the OTU dataset **(Table 2.6)**. However, 273 of the 704 ASVs could only be classified to the taxonomic level of Kingdom. A BLAST [205] query of these sequences against the 16S rRNA gene sequences of bacterial and archaeal type strains indicated that no sequences could be matched with ≥ 86% sequence identity and with a query coverage ≥ 50%. Widening the BLAST query to the entirety of organismal type strains in the Nucleotide collection database revealed that 212/273 (77.7%) ASVs had ≥ 98% sequence match for the *Homo sapiens neanderthalensis* mitochondrial genome. We therefore suspected that these sequences were most likely obtained from human mitochondria, and a BLAST query aligning the 273 unclassified bacteria against the human mitochondrial 16S rRNA gene (GenBank accession # NC_012920) confirmed that these 212 ASVs were ≥ 98.3% matches to human mitochondria. This was in contrast to the primary nested PCR dataset, which contained only 16 unclassified bacterial ASVs, of which two were ≥ 98.9% matches to the human mitochondria 16S rRNA gene. To address the potential influence of mitochondrial sequences on the secondary touchdown PCR dataset analyses, the analyses were repeated after removing the 212 mitochondrial ASVs; the other unclassified bacteria were conservatively left in the dataset. Before subsampling, four placental samples and two technical controls (one from the 5th extraction round) were removed due to sequence counts below 500. After subsampling, all samples had ≥ 97.2% Good's coverage values and there were 430 ASVs retained from the 492 non-mitochondrial ASVs in the dataset. Statistical analyses of this dataset indicated that the mitochondrial ASVs had indeed been the primary drivers of the difference observed between the bacterial profiles of the placental samples, because removing these mitochondrial ASVs eliminated the difference **(Table 2.6, Figure 2.5B & C, Figure 2.10)**.

Table 2.6. Comparison of bacterial community structure of OTU and ASV approaches for the secondary touchdown dataset. ASV and the original OTU datasets were subsampled to 500 (samples were excluded if they had < 500 reads) and only samples ≥ 95% coverage were included in subsequent analyses.

		Bacterial community structure					
Classification method		**OTUs**		**ASVs**		**ASVs after removing human mitochondrial sequences**	
		F	**p-value**	**F**	**p-value**	**F**	**p-value**
Source	*Sample type*	0.767	0.841	2.081	0.001	1.040	0.371
	Round	0.902	0.614	2.377	0.002	1.023	0.405
	Interaction	0.743	0.872	1.901	0.005	0.816	0.855

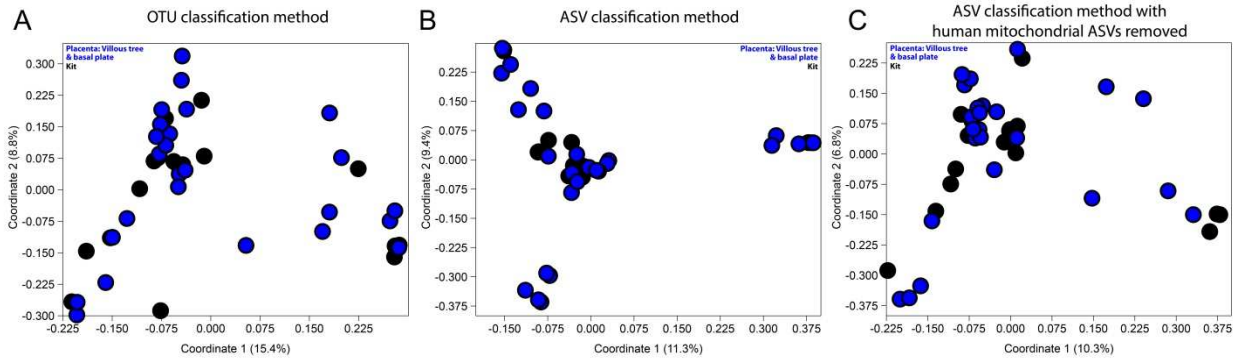

Figure 2.10. Principal Coordinates Analyses (PCoAs) illustrating the structure of the bacterial profiles of placental samples and technical controls from the secondary touchdown PCR dataset under different classification and contaminant removal approaches. Three variants of the secondary touchdown dataset are illustrated: Panel A with the OTU approach, Panel B with the ASV approach, Panel C with the ASV approach after removing ASVs identified as human mitochondria. No differences are seen between Panels A and C, suggesting that mitochondrial sequences are driving any potential variation seen in Panel B. Sample datasets were subsampled to 500 sequence reads (samples were excluded if they had < 500 reads) and only samples with ≥ 95% Good's coverage values were included.

Bacterial profiles of placental tissues from the secondary touchdown PCR dataset after removing any ASVs identified in controls

For the secondary touchdown PCR dataset, following removal of any ASVs that were identified in controls, 121/492 (24.6%) of the non-mitochondrial ASVs were retained. The total read count dropped from 383,121 to 64,501. Average sequence read counts for the placental samples dropped from 13,211 to 2,224, with many losing over half of their reads. Despite losing 83.2% of the total reads from the placental dataset, Good's coverage values for all placental samples > 92%. While the secondary touchdown PCR dataset retained more reads (16.8%) than the primary nested PCR dataset (0.3%) after removing all the ASVs identified in technical controls the same trend of only a modest bacterial signal persisting in placental samples and a lack of consistency in observed bacterial taxonomies across placental samples was again evident. While 7/121 (5.8%) remaining ASVs in the dataset occurred in more than four placental samples (detected in anywhere from 5 to 23 different placental samples, **Figure 2.11**), six of these were subsequently matched to human genes. While any mitochondrial ASVs had been removed prior to this analysis, a BLAST query of these six ASVs against the 16S rRNA Bacteria and Archaea database revealed no matches. However, querying the entirety of the Nucleotide collection database revealed that all six of these ASVs had ≥ 99.4 % similarity to human genes. The other ASV (ASV 217) was identified as *Anaerococcus,* with an average relative abundance of 2.9%. This ASV occurred in seven placental samples, at a relative abundance < 3%, with the exception of one placental sample in which this ASV accounted for 76.0% of the bacterial profile. The majority (15/27) of prominent ASVs (ASVs with an average relative abundance ≥ 1% in placental samples) could each be attributed to an individual sample **(Figure 2.11)** further illustrating the lack of consistency in bacterial signals across placental samples.

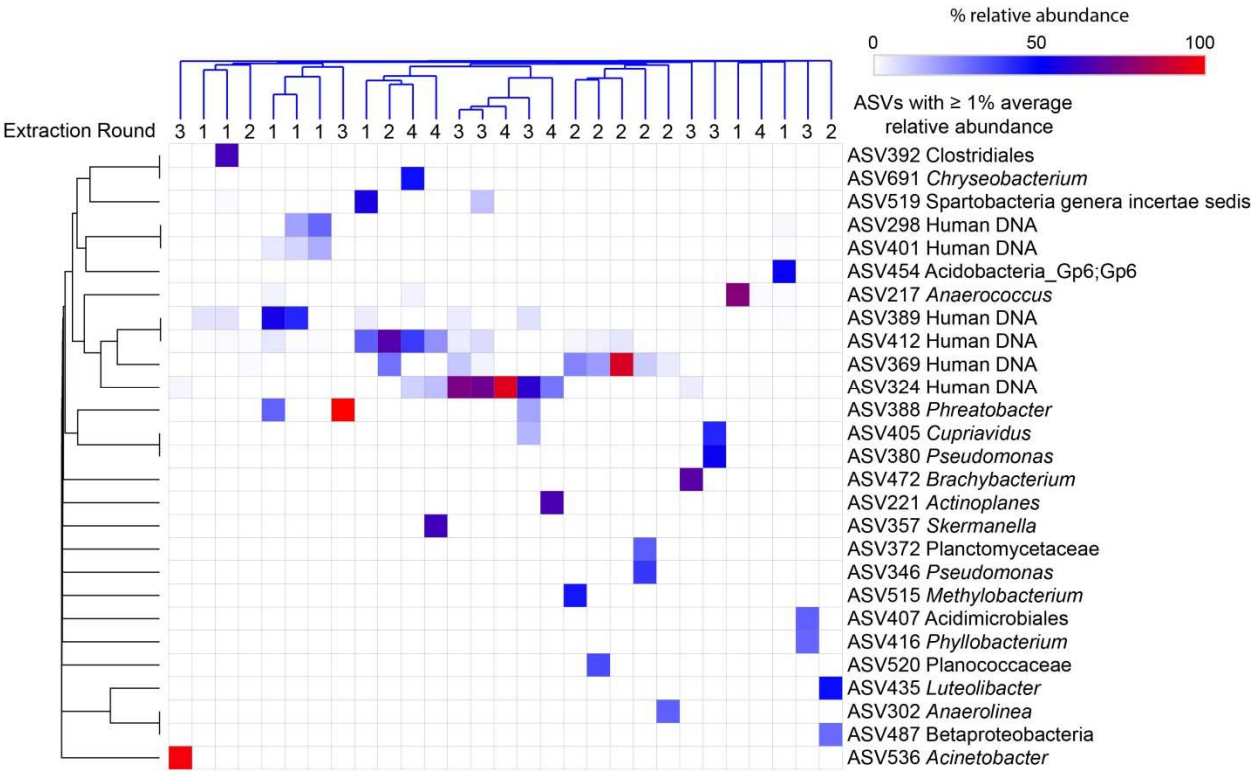

Figure 2.11. Heatmap of the bacterial profiles of placental samples from the secondary touchdown PCR dataset after the removal of all ASVs identified in technical controls. Each column represents a placental sample, and the round of extraction that the sample was processed in is indicated. Samples on the x-axis were clustered based on the Bray-Curtis similarity index. ASVs on the Y-axis were clustered using one minus Pearson's correlation with an average linkage method.

Bacterial profiles of placental tissues from the secondary touchdown PCR dataset after filtering out likely background DNA contaminants identified through the program decontam

The secondary touchdown PCR dataset after *decontam* filtering also mirrored the analysis done without *decontam* filtering for both *isContaminant* and *isNotContaminant* approaches **(Table 2.7, Figure 2.12)**. The exception to this was when using the conservative *decontam* threshold of 0.2. Analyses in which this threshold was used indicated that sample type (i.e. placental sample or technical control) contributed to bacterial profile structure **(Table 2.7)**. This could partly be explained by the profiles of five blank extraction kits being almost entirely (≥ 99.8% of their sequences) composed of a single ASV (ASV8 *Pelomonas aquatica*), and 6/10 kits having > 80.4% their profiles being comprised of this ASV **(Figure 2.13)**. Because this ASV was, in essence, the entire bacterial profile for half of the blank extraction kits, overshadowing the variation in the other blank extraction kits, a subsequent analysis in which those 5 control samples were removed had no differences in bacterial profiles between placental samples and the remaining extraction kit controls (NPMANOVA: Bray-Curtis, sample type, $F = 0.730$, $p = 0.828$). Additionally, this ASV was also the most relatively abundant ASV in placental samples, with an average relative abundance of 15.0%; it was identified in 16/19 (84.2%) placental samples. Only three ASVs occurred in at least nine (47%) placental samples. These ASVs were classified as *Pelomonas aquatica, Ralstonia syzygii*, and a Spartobacteria. In all three cases, these ASVs were also prominent (≥ 1% average relative abundance) in kit controls. The next three most frequently detected ASVs in placental samples were identified as human gene sequences upon BLAST query. Indeed, 6/23 prominent ASVs in placental samples were identified as human gene sequences, and 5/10 ASVs most frequently occurring (in ≥ 5 placental samples) were human-derived. Of the remaining 5 ASVs occurring in ≥ 5 placental samples, three were the previously discussed *Pelomonas aquatica, Ralstonia syzygii*, and Spartobacteria

Table 2.7. Comparison of the structure of the bacterial profiles of placental samples and technical controls for the secondary touchdown PCR dataset after decontam filtering of ASVs. All datasets had human mitochondrial ASVs removed before filtering and analysis. Sample datasets were subsampled to 500 sequences each (samples were excluded if they had < 500 reads) and only samples with ≥ 95% Good's coverage values were included. Non-parametric multivariate analysis of variance (NPMANOVA) based on the Bray-Curtis index was used to assess variation among placental samples and technical controls. Sample type refers to placental sample versus technical control. Round refers to the four different rounds of DNA extractions that were performed; each round of extractions included placental samples and technical controls.

	Bacterial community structure							
	ASVs		***Decontam* filtered**					
			***isContaminant* 0.5 threshold**		***isNotContaminant* 0.2 threshold**		***isNotContaminant* 0.6 threshold**	
Source	**F**	**p-value**	**F**	**p value**	**F**	**p value**	**F**	**p value**
Sample type	1.040	0.371	1.034	0.389	2.045	0.026	1.140	0.250
Round	1.023	0.405	1.108	0.262	1.218	0.220	1.007	0.411
Interaction	0.816	0.855	0.893	0.678	0.663	0.838	0.832	0.717

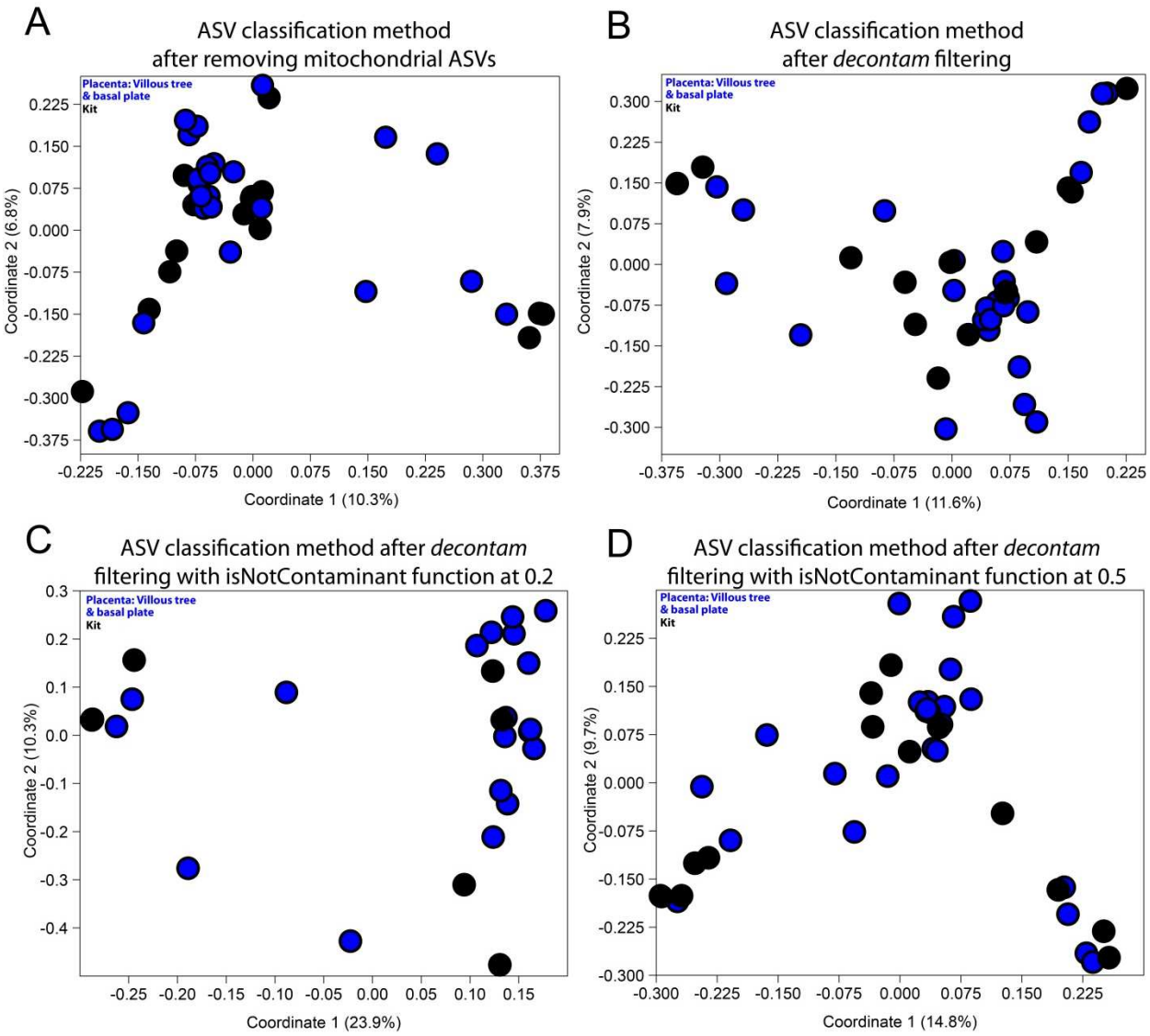

Figure 2.12. Principal Coordinates Analyses (PCoAs) of the structure of the bacterial profiles of placental samples and technical controls from the secondary touchdown PCR dataset after different filtering approaches with the program *decontam*. Panel A is the same as Figure 2.9C. In panel B, ASVs identified as contaminants by the *isContaminant* function in decontam were removed. For panels C and D, ASVs were removed if they were classified by the *isNotContaminant* function in decontam, using a threshold of 0.2 for panel C, and a threshold of 0.5 for panel D. For all panels, sample datasets were subsampled to 500 sequence reads (samples were excluded if they had < 500 reads) and only samples with ≥ 95% Good's coverage values were included. All datasets had human mitochondrial ASVs removed before filtering and analysis.

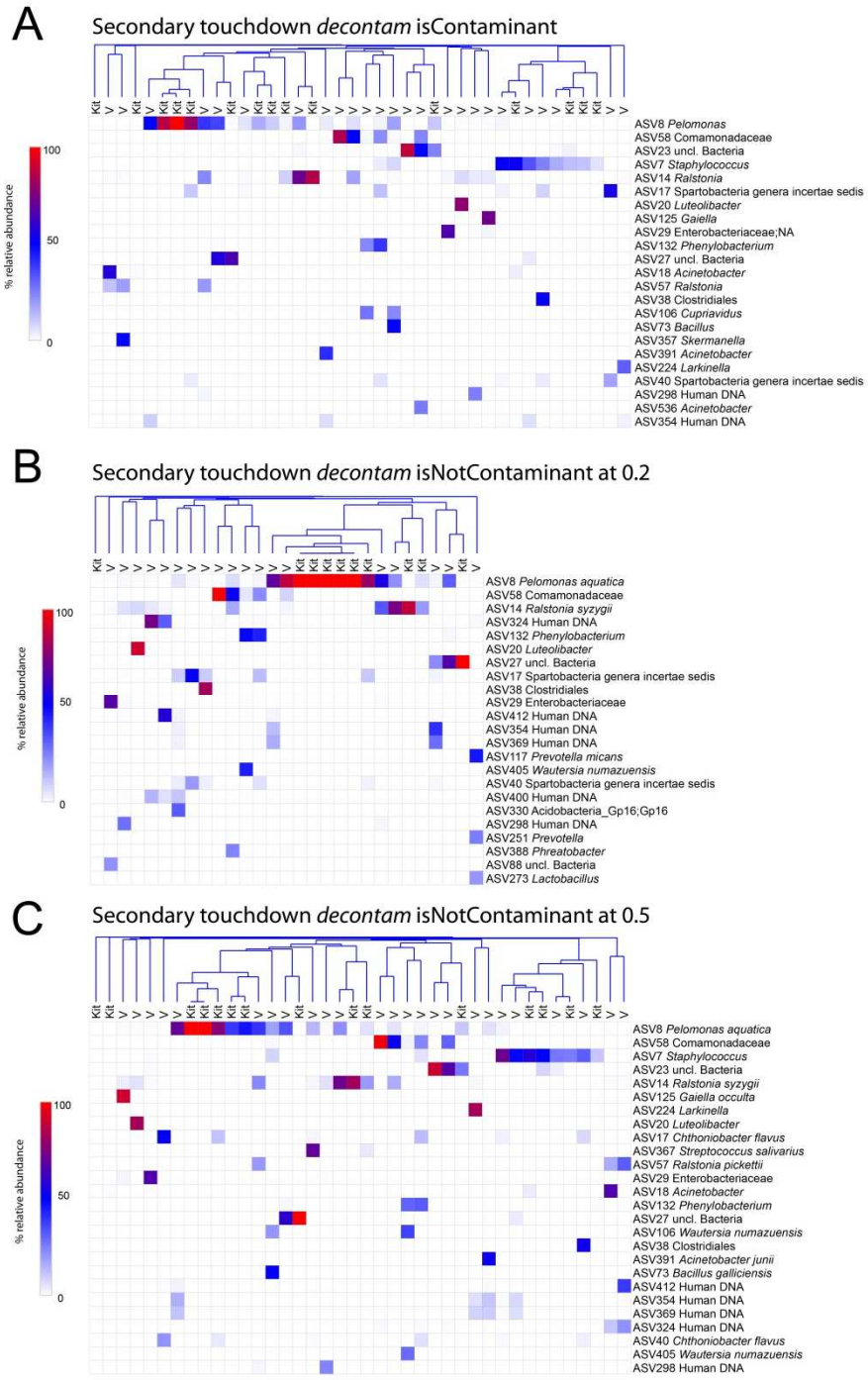

Figure 2.13. Heatmaps of the bacterial profiles of placental samples and technical controls from the secondary touchdown PCR dataset after different filtering approaches with the program *decontam*. Bacterial profiles of villous tree & basal plate (V) samples and background DNA contamination controls cluster indiscriminately, except for panel B wherein several kits are composed almost entirely of a single ASV (*Pelomonas aquatica*). The Y-axis indicates ASVs with ≥ 1% average relative abundance in placental samples for each respective dataset. Samples on the x-axis were clustered based on the Bray-Curtis similarity index.

above and the two other ASVs were a Comamonadaceae and another Spartobacteria. The Comamonadaceae sequence (ASV 58) was 100% identical to that of *Delftia*. *Delftia* is a well-documented contaminant of sequence-based studies and it was here only present in placental samples from a single round of extractions. The sequence associated with this ASV was also identical to that of ASV 634 in the primary nested PCR dataset. In that dataset, ASV 634 was relatively abundant in both placental samples and technical controls. The ASV classified as Spartobacteria (ASV 40) was detected in five placental samples among different extraction rounds, but it was also identified in a kit control. Interestingly, this Spartobacteria was detected in the same samples as the more prominent Spartobacteria ASV (ASV 17), suggesting ASV 40 could be a result of sequencing error. Overall, however, no consistent pattern in bacterial signal is evident among placental samples that distinguishes these samples from technical controls **(Figure 2.13A)**.

V. Discussion

Pregnancy involves an incredibly complex reorganization of physiological and anatomical processes. Not only does the mother have to sustain new life for 9 months, protection and tolerance are necessary to ensure a healthy pregnancy and term gestation. Considering the placenta serves as a critical component to this entire process, complications affecting the placenta are readily detrimental to the fetus and potentially the mother. Thus, comprehensive knowledge of the intrauterine and intra-amniotic environments is critical to prevent and treat these complications. Given that decades of medicine have operated under the understanding that the placenta and intra-amniotic environments are sterile under healthy conditions, concluding that this is not the case should require incontestable evidence. In this study, we found no evidence to support the claim that there is a placental microbiota. We approached the question

with due consideration for the inherent caveats and limitations associated with the various molecular microbiological methods that have been employed previously to conclude there is a placental microbiota [44, 45, 49, 111, 206]. By capitalizing on multiple methodologies, we were able to demonstrate that there is a lack of evidence of a placental microbiota, consistent with the recent research of others [128, 149, 167], and that any shift in paradigms of sterility with respect to the placenta and the fetus would be premature.

Metagenomic surveys of a potential placental microbiota

The metagenomic data presented in this chapter revealed that the bacterial profiles of placental samples as characterized through metagenomics were not consistent with the those from 16S rRNA gene surveys and, although limited in our ability to compare these profiles to those of background technical controls, out of the entire dataset there was only a single bacterial genus not identified in the controls, which was limited to very low relative abundance in a single placental sample. In general, the metagenomic data from placental samples were consistent with DNA contamination in that over one-half of the bacterial sequences were from plant-associated or photosynthetic bacteria. The importance of recognizing the effects of DNA contamination in microbiome studies was highlighted in a recent commentary [100], which emphasized that the data obtained from investigations of low microbial biomass environments should be evaluated through the lens of microbial ecology. In fact, the authors specifically suggest skepticism regarding sequence data that indicate photosynthetic bacteria inhabit internal organs in the human body (e.g. the placenta), which would prohibit photosynthesis. Furthermore, it is unlikely that microorganisms associated with non-human, non-mammalian environments, such as the plant pathogens “Candidatus Phytoplasma”, *Xanthomonas,* and *Xyella* , the aquatic bacteria

Beggiatoa, Roseobacter, Hahella, and *Halangium*, or the algal symbiont *Dinoroseobacter*, could readily grow and reproduce within the ecologic niche of the placenta.

In thorough consideration of a placental microbiota, our metagenomic data did identify sequences from ecologically-plausible microorganisms warranting further discussion. *Coprobacillus* was detected in placental samples and constituted 30.5% of the bacterial sequences. This genus has been detected in two sequence-based studies of term and preterm placentas at low abundance [45, 207], however, it was not detected in any of our 16S rRNA gene surveys. Although the primers used to target the 16S rRNA gene in the first round of amplification in the primary and secondary nested PCRs (27F/1492R; 341F/1061R) were not an exact match for *C. cateniformis*, the only member of the genus *Coprobacillus* [208], the primers used for the secondary standard and touchdown PCRs (515F/806R) were a perfect match for this bacterium. Therefore, if *Coprobacillus* was present in placental tissues and if its 16S rRNA gene sequence was similar to that of the lone characterized representative of this genus, we should have identified it in the 16S rRNA gene analyses.

Streptomyces is another microorganism to consider as it was, on average, 1% of the bacterial sequences in the metagenomic data obtained from placental tissues and has been previously detected in sequencing studies of placentas at term [44, 45, 84]. In the 16S rRNA gene surveys of our study only 2 sequences were assigned to *Streptomyces*, and for this genus, the V4 primers (515F/806R) were perfect matches for nearly all (98.6%) of the 588 type strains found in the Ribosomal Database Project [209]. This suggests that if *Streptomyces* were actually present in these placental samples, it should have exhibited a stronger signal than two sequences in the standard 16S rRNA gene PCR and touchdown PCR analyses. Although even less abundant than *Streptomyces* (>0.1% average relative abundance), other bacterial genera identified in the

metagenomic data of our study and also in other sequence-based studies of placental tissues at term include *Neisseria* [44, 45], *Rhodococcus* [44, 111], *Clostridium* [210], *Streptococcus* [49, 95], and *Burkholderia* [45]. Again however, in our study, sequences for these microorganisms were detected in all placental samples and background technical control samples and all five have been identified as contaminant sequences in prior sequence-based studies [99, 119, 211]. Therefore, both alone, and more importantly in consideration with the other methodologies, we did not find sufficient evidence to conclude that the bacterial signals identified through the metagenomic sequencing represent evidence of a placental microbiota and, given the diminutive fraction of reads attributed to functional genes, a functional bacterial ecosystem in this human organ.

Classifying 16S rRNA gene sequence data as operational taxonomic units (OTUs) versus as amplicon sequence variants (ASVs)

The caveats and limitations associated with low microbial biomass studies, especially studies of the placenta, have spawned attention in the field [97-99, 119, 149], including the recent review by O'Callaghan et al. that contained a reanalysis of our study's data [103]. While they did not reach a different conclusion than us regarding the existence of a placenta microbiota, they did suggest there were some noteworthy results that differed from our original OTU analysis presented in Theis et al. [150]. However, we identified specific issues with the review's analysis. First, samples were grouped together for sequence read counts and heatmaps, which inflates the magnitude of the signals. Second, their visualization of β-diversity was limited to eight out of the total 58 placental samples included in our original study, without explanation. One potential explanation could be that subsampling led to the exclusion of the majority of samples, but this was not communicated by the authors. Third, they assert that the pooled placental samples had > 200,000 sequence reads, even after removing the ASVs identified in

technical controls, yet in our reanalysis, the result is quite different [maximum read count for any sample was 2206 (not shown), and the average read counts were much lower; **Figure 2.6**]. Fourth, the review documented the presence of an ASV classified as *Ureaplasma* in the room control samples that is completely discordant with our own analyses. Specifically, the review pooled control sample types and indicated in the heatmap that *Ureaplasma parvum* had a read abundance ≥ 3,000 in room controls. However, in our original OTU analysis, individual room controls contained no more than two reads for *Ureaplasma* OTUs and no more than one read after reclassifying the sequences through the DADA2 ASV pipeline, prior to any subsampling. To help illustrate this in the Results section of this chapter, *Ureaplasma* ASVs were included in the heatmaps even if they did not meet the cutoff for inclusion as a prominent ASV in a given analysis **(Figure 2.5A, Figure 2.9)**.

Although the review provided very limited details regarding how our data were re-analyzed, restricting our ability to replicate the exact analyses done in the review, our reanalysis was conducted with the methods we have ascertained to be the most appropriate and comprehensive for this type of microbiome data. We performed the OTU/ASV reanalysis on the primary nested PCR dataset and the secondary touchdown PCR dataset from Theis et al. [150], as these were the two PCR amplification methods that yielded large numbers of quality sequences and yet were not dominated almost exclusively by *Escherichia coli* sequences, as was the case with the secondary nested PCR dataset (as discussed in Theis et. al. [150]). With each of these two datasets, we performed three different analyses to address differences between OTU and ASV classification approaches, as well as the implementation of the program *decontam* to identify and remove likely background DNA contaminants. First, we directly compared the original OTU dataset with a dataset in which the sequences were reclassified using ASVs. In

doing so, we subsampled both datasets to a sequence depth of 500 (a depth of 500 sequences was also used by O'Callaghan et al. [103]). Additionally, all our analyses were limited to samples with ≥ 95% Good's coverage, because samples with lower Good's coverage values are unlikely to provide thorough representations of the actual diversity in those samples; Good's coverage was not addressed in the review. Second, we analyzed the ASV dataset after removing all ASVs detected in controls without any subsampling (an approach also featured in the review [103]). Third, we utilized the program *decontam* [120] to identify and remove any ASVs determined to be contaminants from the ASV dataset. This analysis was performed two ways: following methods described by O'Callaghan et al. [103] using the *isContaminant* function, and using the *isNotContaminant* function, which we believe to be the more appropriate way to use the *decontam* tool, in the context of low microbial biomass samples, as suggested by Davis et. al. [120]. Upon re-analysis of our own study after ASV classification and with detailed methods to account for our results and conclusions, despite some subtle differences to our original work, we validated our previous findings and conclusions; there is a lack of evidence of a human placental microbiota from placentas delivered at term.

As technologies improve and new computational tools are explored, the sequencing and characterization of presumed sterile environments will remain a target of investigation for microbial signals. Because sequencing surveys produce enormous amounts of data, the way these data are processed and analyzed can have a significant influence over observed results and subsequent conclusions. Efforts have been made to create standardized and reproducible ways to analyze these data, however, preferences among researchers will always exist. In the case of 16S rRNA gene sequencing, OTUs have been the most common approach for classifying sequences. While the microbiome community is beginning to embrace ASVs as a valuable tool, suggesting

they should replace OTUs as the predominant method to classify 16S rRNA gene sequences may not be appropriate in all circumstances. The thorough reanalysis of the data from our placental study [150] demonstrates that for these two methods, we found very little difference. It is, however, critical to consider the context. In this case, the environment being investigated in our study is one classically considered sterile. The advantages that ASV classification have over OTU classification will likely be limited to environments with a substantive microbial community. In the context of low microbial biomass, sequencing technologies are producing a larger proportion of sequences from background DNA contaminants, which are more likely to be of lower quality and consequently more prone to sequencing errors. By analyzing 16S rRNA gene data of this sort as ASVs, the single nucleotide resolution is likely to detect an increased number of ASVs, an artifact of sequencing the poor quality background DNA that is likely in greater abundance in low microbial biomass than in high microbial biomass environments. Analysis of ASV data from a situation like this may lead to an observation of strain level variation of microbial communities that does not actually exist and thusly, multiple approaches to the data are warranted (e.g. analysis with multiple tools/methods). This example further illustrates the caveats associated with studies investigating environments with low microbial biomass and the importance of robust methodologies and analysis.

CHAPTER 3: DOES THE MOUSE PLACENTA HAVE A MICROBIOTA? CULTURE AND MOLECULAR SURVEYS OF THE MURINE MICROBIOTA

The data presented in this chapter are part of the manuscript "No Consistent Evidence for Microbiota in Murine Placental and Fetal Tissues," published January 2020 in *mSphere* by Theis, K.R., Romero, R., **Greenberg, J.M.**, et al. [212], of which I am third author behind two senior authors. Some of the text of this Chapter (i.e. Methods and Results) appears as it does in the published manuscript. Dr. Andrew Winters generated the quantitative real-time PCR data for this study. Madison Ahmad, a Master's student in our laboratory, contributed to the bacterial culture component of this study and was responsible for the plate wash PCR and 16S rRNA gene sequencing of plate washed samples.

I. Abstract

The existence of a placental microbiota and *in utero* colonization of the fetus have been the subjects of recent debate. While the bulk of this work has been focused on humans, the mouse model presents a unique opportunity for more in-depth investigation of mammalian pregnancy. In this chapter, my objective was to complement our previous work on the human placenta by evaluating whether the placental and fetal tissues of mice harbor bacterial communities using multiple methodologies with comparisons to maternal samples and background technical controls. Bacterial profiles of the placenta and fetal brain, lung, liver, and intestine samples were characterized through culture, quantitative real-time PCR (qPCR), and 16S rRNA gene sequencing. Maternal samples included the mouth, lung, liver, uterus, cervix, vagina, and intestine. Positive bacterial cultures from placental and fetal tissues were rare; of the 165 total bacterial cultures of placental tissues from the 11 mice included in this study, only nine yielded at least a single colony, and five of those nine positive cultures came from a single mouse. Cultures of fetal intestinal tissues yielded just a single bacterial isolate: *Staphylococcus hominis*, a common skin bacterium. Bacterial loads of placental and fetal brain, lung, liver, and intestinal tissues were not higher than those of DNA contamination controls and did not yield substantive 16S rRNA gene sequencing libraries. Overall, from all placental or fetal tissues, there

was only a single bacterial isolate that came from a fetal brain sample having a bacterial load higher than that of contamination controls and that was identified in sequence-based surveys of at least one of its corresponding maternal samples. Therefore, using multiple modes of microbiologic inquiry, there was not consistent evidence of bacterial communities in the placental and fetal tissues of mice.

II. Introduction

As previously discussed in the Introduction and Chapter 2 of this thesis, the existence of a placental microbiota has been the focus of debate for over half a decade. Utilization of molecular surveys has led some to the conclusion that a placental microbiota exists in healthy pregnancies; however, as shown in Chapter 2, the caveats associated with these surveys must be thoroughly complemented through alternative methodologies. Considering the well-documented relationships between hosts and their microbiota, a placental microbiota would likely have been conserved among mammals (i.e. if humans have a placental microbiota then other mammals likely do as well). Several studies have investigated the placenta for bacterial communities in mice and rats [83, 213, 214], and as with human placental studies [44, 45, 49, 111, 206, 215], they have concluded that there are bacterial communities in these tissues. Investigation of a placental microbiota in animal models has the advantage of being able to surgically remove the placenta and fetal tissues before the onset of labor, a process that can introduce microbes into the upper reproductive tract [216-218]. Additionally, if *in utero* colonization is occurring, the source of the colonizing bacteria must be the mother, and with animal models, researchers can collect an array of samples from the mother for investigation and contrast that are not obtainable from human subjects.

After our evaluation of a placental microbiota in humans and the subsequent lack of evidence for it [150] presented in Chapter 2, I transition from the human model to the mouse model for which we can validate our findings from the human placenta and further support these findings with additional fetal and maternal tissues. As established in Theis et al. [150], the evidence necessary in establishing that there exists bacterial communities in placental or fetal tissues requires: 1) the identification of bacterial DNA in placental or fetal tissues that is distinct from bacterial DNA detected in technical controls (e.g. DNA extraction kits, PCR reagents, laboratory environments), 2) confirmation that the bacterial load of placental or fetal tissues exceeds that of technical controls through quantitative real-time PCR (qPCR), 3) visualization of bacteria in placental or fetal tissues using microscopy, 4) demonstration of the viability of bacteria in these tissues through culture, and 5) ecological plausibility (i.e. the detected bacteria could survive and reproduce in these tissues) [100]. To date, these criteria have not been met in any one study. In the previous chapter, we evaluated the human placenta thoroughly addressing the first, second, fourth, and fifth criteria and were unable to find evidence supporting the existence of a placental microbiota. The objective of this chapter is to validate those findings by thorough evaluation of placental and fetal tissues from mice, similar to our human study, via bacterial culture, qPCR, and 16S rRNA gene sequencing, and by comparing them to background technical controls and maternal tissues for additional contrast.

III. Methods and Materials

Study subjects and sample collection

Eleven pregnant C57BL/6 (B6) mice were anesthetized with isoflurane at 17.5 days gestation. The dam's abdomen was shaved, and alcohol was liberally applied to the abdomen. Maternal blood was collected by cardiac puncture and death was assured by cervical dislocation.

The dam was then placed on a surgical platform within a biological safety cabinet. Study personnel (i.e. myself or Dr. Theis) donned sterile surgical gowns, masks, full hoods, and powder-free exam gloves during sample collection. The oral cavity and vagina were swabbed with Dacron and ESwabs for molecular microbiology and bacterial culture, respectively. For the abdomen, a Dacron swab was collected, iodine was applied and, after the iodine dried, an ESwab was collected. A midline incision was made along the full length of the abdomen. The peritoneum was sampled with a Dacron swab. The uterine horns were separated from the cervix and placed within a sterile petri dish, wherein they were immediately processed by a different investigator within the biological safety cabinet. Uterine horns were dissected and fetuses (the fetus inside the amniotic sac attached to the placenta) were placed in individual Petri dishes. Uterine tissues were collected for both molecular microbiology and bacterial culture. Two fetuses from each dam were selected for analysis; tissues from one were used for molecular microbiology and tissues from the other for bacterial culture. From each fetus, the placenta, lung, liver, intestine, and brain (molecular microbiology was performed on fetal brain samples from all 11 mice; bacterial culture was completed on fetal brain samples from mice E-K) were collected. The fetal spleen and tail were also collected for molecular microbiology.

Next, the maternal cervix, liver, and lung were sectioned and one sample of each was placed into a sterile 1.5 ml microcentrifuge tube and an anaerobic transport medium tube for molecular microbiology and bacterial culture, respectively. Lastly, after all placental and fetal tissues were sampled and stored, the maternal heart and the maternal proximal and distal intestine were collected for molecular microbiology, and the maternal middle intestine was collected for bacterial culture. Procedures were approved by the Institutional Animal Care and Use Committee at Wayne State University (Protocol 18-03-0584).

Mouse tissue processing

Maternal and fetal mouse tissue samples within anaerobic transport containers or ESwabs were delivered to the lab immediately after surgery was completed on the final mouse of the day. Upon receipt into the lab, samples were brought into a hypoxic chamber under 5% CO_2, 5% O_2, 90% N_2 atmospheric conditions and processed in the following order processed in the following order: placenta, fetal liver, fetal lung, fetal brain, fetal intestine, maternal uterus, maternal liver, maternal lung, maternal cervix, maternal skin post-sterilization, maternal vagina, maternal oral cavity, and the maternal mid-intestine. While processing samples for bacterial culture within the chamber, study personnel wore sterile sleeve protectors, nitrile exam gloves, and sterile nitrile gloves over the top of the nitrile exam gloves. Tissues were removed from anaerobic transport medium tubes using a sterile disposable inoculating loop, placed into a dounce reservoir (2ml or 5 ml) containing 1ml of sterile PBS, and carefully homogenized for one minute. The tissue homogenates were then transferred into a 5 ml tube containing 1.5 ml of sterile PBS. Maternal lung and maternal mid-intestine tissues for mice E-K were homogenized in sterile 5ml tubes using 0.5ml PBS and a sterile disposable scalpel. Each ESwab sample was vortexed thoroughly and the container's medium was transferred into a 5ml tube containing 1.5ml of sterile PBS using sterile disposable transfer pipettes. After the last tissue was processed for a mouse, each tissue homogenate was split between 3 sterile tubes for culturing under three atmospheric conditions.

During the processing of each mouse's samples for culture, three chocolate agar plates were left in the hypoxic chamber to serve as negative controls; they were subsequently incubated for seven days under oxic, hypoxic, and anoxic conditions. Additionally, for each mouse the PBS stock used for tissue homogenization was plated on blood agar, chocolate agar, and MacConkey

agar, and was further added to SP4 broth containing urea and arginine. The PBS-control blood agar and chocolate agar plates were incubated under all three atmospheric conditions, and the MacConkey agar plates and SP4 brothers were incubated under oxic conditions only. These negative controls were incubated for seven days.

Mouse tissue cultivation methods

Tissue homogenates and swab solutions were cultured for 7 days under three atmospheric conditions: anoxic (5% CO_2, 10% H_2, 85% N_2), hypoxic (5% CO_2, 5% O_2, 90% N_2), and oxic (ambient). Every tissue was plated on tryptic soy agar with 5% sheep's blood and chocolate agar plates in duplicate under each atmospheric condition; under oxic conditions all tissues were also plated on MacConkey's agar, also in duplicate. For each plate, 100ul of tissue or swab homogenate was pipetted onto the center and thoroughly spread over the plate until dry using an L-shaped spreader. 100ul of each fetal tissue and maternal reproductive tissue were inoculated in SP4 broth with urea and SP4 broth with arginine. For each atmosphere and media type a PBS control was included following the same protocol as above by inoculating 100ul of sterile PBS.

Agar plates and SP4 broth tubes were observed for days 1-4 and day 7 (each day represents ~24 hours after plating). Colonies were counted when appropriate (< ~25 colonies) and any colonies observed on placental tissues, fetal tissues, and negative control plates were restreaked for isolation and purity. Pure colonies from restreaked plates were saved in 2mL cryovial tubes containing PBS with 15% glycerol, in triplicate when possible, and stored at -80 □. An additional colony was saved in a 1.5ml tube containing ≤ 500ul of sterile PCR-grade H_2O for downstream taxonomic identification via Sanger sequencing of the 16S rRNA gene (described below). Contiguous growth was observed on a single plate for one placental sample

(mouse J); the isolates all had the same morphotype, so only representative isolate was streaked for purity and downstream taxonomic identification.

The negative-control plates yielded five total bacterial isolates over the course of the experiment. Four were successfully sequenced: two were identified as *Cutibacterium acnes*, and two were identified as *Staphylococcus hominis*. If a specific bacterium was cultured on a technical control plate on the day a mouse's samples were processed as well as on a placental or fetal sample plate for that mouse (i.e., there was a 100% 16S rRNA gene sequence match between the bacterial isolates recovered on the two plates), that bacterium was not included in analyses. Overall, this included 11 bacterial isolates for three mice (D, J, and K). Of these 11 isolates, four were *C. acnes*, and seven were *S. hominis*. If a specific bacterium was cultured on a mouse's placental or fetal sample plate as well as on a technical control plate from another sample processing day, but not on a control plate from that mouse's sampling day, the bacterium was included in analyses.

For maternal cervix, uterus, and liver samples, the unique isolate morphotypes on each plate were streaked for purity and taxonomically identified through Sanger sequencing of the 16S rRNA gene. Maternal samples from the intestine, vagina, mouth, and lung and any other plates with contiguous growth (either too many colonies to count or a lawn of growth) were saved for downstream plate wash PCR and 16S rRNA gene sequencing (detailed below).

Taxonomic identification of individual bacterial isolates

After the bacterial isolates were streaked for purity, the isolates from placental, fetal, and maternal uterine, cervical, and liver samples were stored in nuclease-free water and frozen at -20 □ until colony PCR targeting the 16S rRNA gene was performed. The 16S rRNA gene of each isolate was first amplified using the 27F/1492R primer set and then bi-directionally Sanger

sequenced through GeneWiz using the 515F/806R primer set, which targets the V4 hypervariable region of the 16S rRNA gene. Forward and reverse reads were trimmed using DNA Baser software (http://www.dnabaser.com/) with default settings and assembled using the CAP (contig assembly program) of BioEdit software (v7.0.5.3), also with default settings. The taxonomic identities of individual bacterial isolates were determined using the Basic Local Alignment Search Tool (BLAST) [205]. 16S rRNA gene sequence similarities between isolates and their top match on BLAST were ≥ 99.5%, unless otherwise noted **(Table 3.1,** Table 3.2**)**.

DNA extraction from plate washes of cultured bacteria

Plate wash was performed by pipetting 1-2 ml of PBS onto the agar plate and dislodging bacterial colonies with either sterile L-shaped spreaders or inoculating loops. The PBS wash was then transferred into cryovials and stored at -80°C until DNA was extracted. DNA was extracted from plate wash samples using Qiagen DNeasy PowerSoil extraction kits. Washes from maternal samples that yielded growth under multiple atmospheres for the same media type were pooled prior to the extraction process. Purified DNA was stored at -20 □.

16S rRNA gene sequencing of plate wash extracts

The 16S rRNA genes in plate wash extracts were sequenced at Wayne State University on an Illumina MiSeq system using a 2 X 250 cycle V2 kit, and following Illumina sequencing protocols [219]. The 515F/806R primer set was used to target the V4 region of the 16S rRNA gene. The 16S rRNA gene sequences from the paired fastq files for these samples were processed as previously described [150].

DNA extraction from swab and tissue samples

All Dacron swab and tissue samples collected for molecular microbiology were stored at -80 □ until DNA extractions were performed. DNA extractions were performed in a biological safety cabinet by study personnel donning sterile surgical gowns, masks, full hoods, and powder-free exam gloves. Extractions of tissues generally included 0.015 – 0.100 grams of tissue, except for the fetal tail and spleen, whose masses were very low.

DNA was extracted from swabs, tissues, and background technical controls (i.e. sterile Dacron swabs (N = 11) and blank DNA extraction kits (N = 23)) using the DNeasy PowerLyzer PowerSoil Kit (Qiagen, Germantown, MD) with minor modifications to the manufacturer's protocol. Specifically, 400 µl of bead solution, 200 µl of phenol:chloroform:isoamyl alcohol (pH 7–8), and 60 µl of Solution C1 were added to the supplied bead tube. Cells within samples were lysed by mechanical disruption for 30 seconds using a bead beater. After centrifugation, the supernatants were transferred to new tubes, and 100 µl of solution C2, 100 µl of solution C3, and one µl of RNase A enzyme were added, and tubes were incubated at 4° C for five minutes. After centrifugation, the supernatants were transferred to new tubes that contained 650 µl of solution C4 and 650 µl of 100% ethanol. The lysates were loaded onto filter columns, centrifuged for one minute, and the flow-through was discarded. This step was repeated until all sample lysates were spun through the filter columns. Five hundred µl of solution C5 were added to the filter columns, centrifuged for one minute, the flow-through was discarded, and the tube was centrifuged for an additional three minutes as a dry-spin. Finally, 60 µl of solution C6 were placed on the filter column and incubated for five minutes before centrifuging for 30 seconds to elute the extracted DNA. Purified DNA was stored at -20° C.

Purified DNA was quantified using a Qubit 3.0 fluorometer with a Qubit dsDNA BR Assay kit (Life Technologies, Carlsbad, CA), according to the manufacturer's protocol. All

purified DNA samples were then normalized to 80 ng/µl (when possible) by diluting each sample with the Qiagen elution buffer (Solution C6).

16S rRNA gene quantitative real-time PCR (qPCR)

A preliminary test was performed to investigate whether DNA amplification inhibition existed among the different sample types. For this test, 4.7 µl of purified *Escherichia coli* ATCC 25922 (GenBank accession: CP009072) genomic DNA (0.005 ng/µl) containing seven 16S rDNA copies per genome was spiked into 7.0 µl of purified DNA from mouse samples that were serially diluted with Solution C6 by a factor of 1:3 (i.e. 1:0, 1:3, 1:9). For tissue sample types with a mean DNA concentration above 250 ng/µl, DNA concentrations were normalized to 80 ng/µl by dilution with Solution C6 before being serially diluted and spiked with *E. coli* genomic DNA. Genomic DNA was quantified using a Qubit 3.0 fluorometer with a Qubit dsDNA HS Assay kit (Life Technologies, Carlsbad, CA) according to the manufacturer's protocol. Three µl of each spiked sample were then used as a template for qPCR. For all samples, spiked reactions contained approximately 1.0 x 10^3 *E. coli* 16S rDNA copies. There was no evidence of DNA amplification inhibition (data not shown).

Total bacterial DNA abundance within samples was measured via amplification of the V1 - V2 region of the 16S rRNA gene according to the protocol of Dickson et al [70] with minor modifications. These modifications included the use of a degenerative forward primer (27f-CM: 5'-AGA GTT TGA TCM TGG CTC AG-3') [123] and a degenerate probe containing locked nucleic acids (+) (BSR65/17: 5'-56FAM-TAA +YA+C ATG +CA+A GT+C GA-BHQ1-3'). Each 20 µl reaction contained 0.6 µM of 27f-CM primer, 0.6 µM of 357R primer (5'-CTG CTG CCT YCC GTA G-3'), 0.25 µM of BSR65/17 probe, 10.0 µl of 2X TaqMan Environmental Master Mix 2.0 (Life Technologies, Carlsbad, CA), and 3.0 µl of either purified DNA (diluted to

80 ng/µl when possible), elution buffer, or nuclease-free water. The total bacterial DNA qPCR was performed using the following conditions: 95° C for 10 min, followed by 45 cycles of 94° C for 30 sec, 50° C for 30 sec, and 72° C for 30 sec. Duplicate reactions were run for all samples. All samples were run across a total of five runs.

Raw amplification data were normalized to the ROX passive reference dye and analyzed using the on-line platform Thermo Fisher Cloud: Standard Curve (SR) 3.3.0-SR2-build15 with automatic threshold and baseline settings. Cycle of quantification (Cq) values were calculated for samples based on the mean number of cycles required for normalized fluorescence to exponentially increase.

After plotting a regression of log(*E. coli* 16S rRNA gene copy number) and Cq value for standard curves included in each qPCR run, 16S rRNA gene copy number in mouse samples was calculated according to Gallup [220] using the equation $X_o = E_{AMP}^{(b-Cq)}$, where E_{AMP} is the exponential amplification value for the qPCR assay, calculated as $E_{AMP} = 10^{(-1/m)}$ and b and m are the intercept and slope of the regression.

16S rRNA gene sequencing of swab and tissue sample extracts

Amplification and sequencing of the V4 region of the 16S rRNA gene was performed at the University of Michigan's Center for Microbial Systems as previously described [109], except that library builds were performed in triplicate and pooled for each individual sample prior to the equimolar pooling of all sample libraries for multiplex sequencing. Sample-specific MiSeq run files have been deposited on the NCBI Sequence Read Archive (BioProject ID SUB6641162).

Raw sequence reads were processed using mothur software (v1.39.5) [41] following the Standard Operating Procedure provided by Schloss et al. (www.mothur.org/wiki/MiSeq_SOP).

Paired-end reads were assembled into contiguous sequences, quality checked (maximum length = 275, maximum ambiguous base pairs = 0, and maximum number of homopolymers = 8), and aligned against the SILVA 16S rDNA reference database (release 102) [221, 222]; sequences falling outside the target alignment space were removed. Quality sequences were pre-clustered (diffs = 2) and chimeric sequences were identified with VSEARCH [223] and removed. The remaining sequences were taxonomically classified using the SILVA reference database with a k-nearest neighbor approach and a confidence threshold of 80%. Sequences derived from an unknown domain, Eukaryota, Archaea, chloroplasts, or mitochondria were removed. Operational taxonomic units (OTUs) were defined by clustering sequences at a 97% sequence similarity cutoff using the average neighbor method.

Statistical analysis

The bacterial loads, as assessed through qPCR, of maternal, placental and fetal samples were compared to those of background technical controls (i.e. sterile Dacron swabs and blank DNA extraction kits) using t-tests or Mann-Whitney U tests with sequential Bonferroni corrections applied. The bacterial loads of placental and fetal tissues were compared to one another using Wilcoxon matched pairs tests, again corrected for multiple comparisons.

The beta diversity of 16S rRNA gene profiles among maternal, placental, fetal and technical control samples were characterized using the Bray-Curtis similarity index. Bray-Curtis similarities in sample profiles were visualized using Principal Coordinates Analysis (PCoA) plots and statistically evaluated using non-parametric multivariate ANOVA (NPMANOVA). These analyses were limited to samples that yielded a 16S rRNA gene library with ≥ 250 quality-filtered sequences and a Good's coverage ≥ 95%. All data analysis was completed in PAST

software (v 3.25) [202]. Heat maps of sample bacterial profiles were generated using the open-source software program Morpheus (https://software.broadinstitute.org/morpheus) [204].

IV. Results

Bacterial culture from placental and fetal tissues

Growth of bacterial isolates from placental and fetal tissues was rare **(Figure 3.1A, Figure 3.2).** Only 3/11 mice (F, H & J) yielded more than two total bacterial isolates across all their cultured placental and fetal samples under all growth conditions **(Table 3.1)**. Most of the bacterial isolates from placental and fetal samples were *Staphylococcus* spp. (mostly *S. hominis*) **(Figure 3.1A)**. *Staphylococcus* spp. were cultured from the mouth, intestine, and vagina of mouse dams **(Figure 3.1B)**; however, two of the five bacterial isolates recovered from the 114 negative control plates included in this study were also *Staphylococcus*, specifically *S. hominis*. The non-staphylococci bacteria cultured from placental or fetal samples were *Bacillus*, *Corynebacterium*, *Paenibacillus*, *Propionibacterium*, and unclassified bacilli **(Table 3.1)**. These bacteria were rarely, if ever, cultured from maternal samples **(Figure 3.1A & B)**.

In general, only one or two placental or fetal sites within a given fetus yielded a bacterial isolate, and there was little consistency among the fetuses in terms of which site yielded an isolate **(Table 3.1, Figure 3.1A)**. For example, of the 132 blood and chocolate agar plates on which placental tissue homogenates were spread, only nine (6.8%) yielded even a single bacterial isolate, and five of these plates came from a single placental sample (Mouse H) **(Figure 3.2)**. All of the bacterial isolates from Mouse H's placental sample were *Staphylococcus* (either *S. hominis* or *S. epidermidis* / *caprae* / *capitis*). There were no exact matches of the 16S rRNA genes of these isolates within the 16S rRNA gene surveys of placental tissues from Mouse H, nor were there any matches within the 16S rRNA gene surveys of any of the sampled maternal body

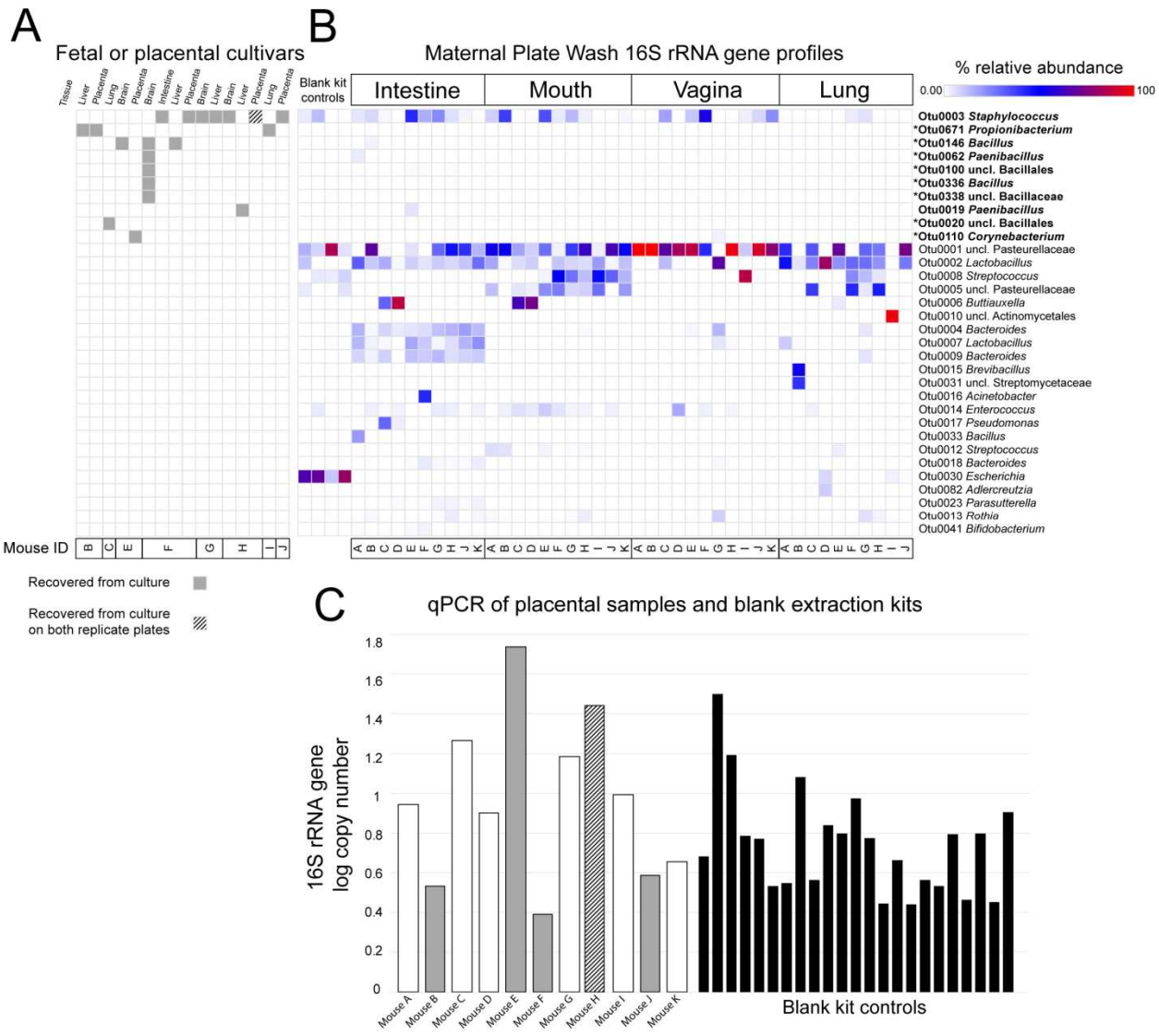

Figure 3.1. Bacterial cultivation results for A) placental and fetal tissues in relation to those for B) maternal intestinal, mouth, vaginal, and lung samples, and C) a comparison of the bacterial loads of individual placental samples and blank extraction kit controls in light of the cultivation results. Panel A indicates the recovery of any bacterial isolates from placenta and/or fetal tissues, by mouse and across different growth media and atmosphere conditions. The taxonomic assignments of these isolates were determined by comparing their 16S rRNA gene sequences to those of the operational taxonomic units (OTUs) of molecular surveys of the mixed bacterial communities cultured from maternal intestinal, oral, vaginal, and lung samples (sequence identity was ≥ 97.2%). Panel B provides the results of 16S rRNA gene molecular surveys of the plate washes of bacterial growth from maternal intestinal, oral, vaginal, and lung samples, as well as of blank extraction kit controls processed alongside the plate washes. OTUs were included in the heat map in Panel B if they had an average percent relative abundance ≥ 0.5% across all plate washes or if they were the best 16S rRNA gene sequence match to bacterial isolates in Panel A (indicated by an asterisk). The bolded OTUs represent the best 16S rRNA gene sequence matches to placental and fetal isolates in Panel A. Panel C illustrates similarities

in bacterial load, as assessed by 16S rRNA gene quantitative real-time PCR (qPCR), between placental samples yielding at least one bacterial isolate and blank DNA extraction kit controls.

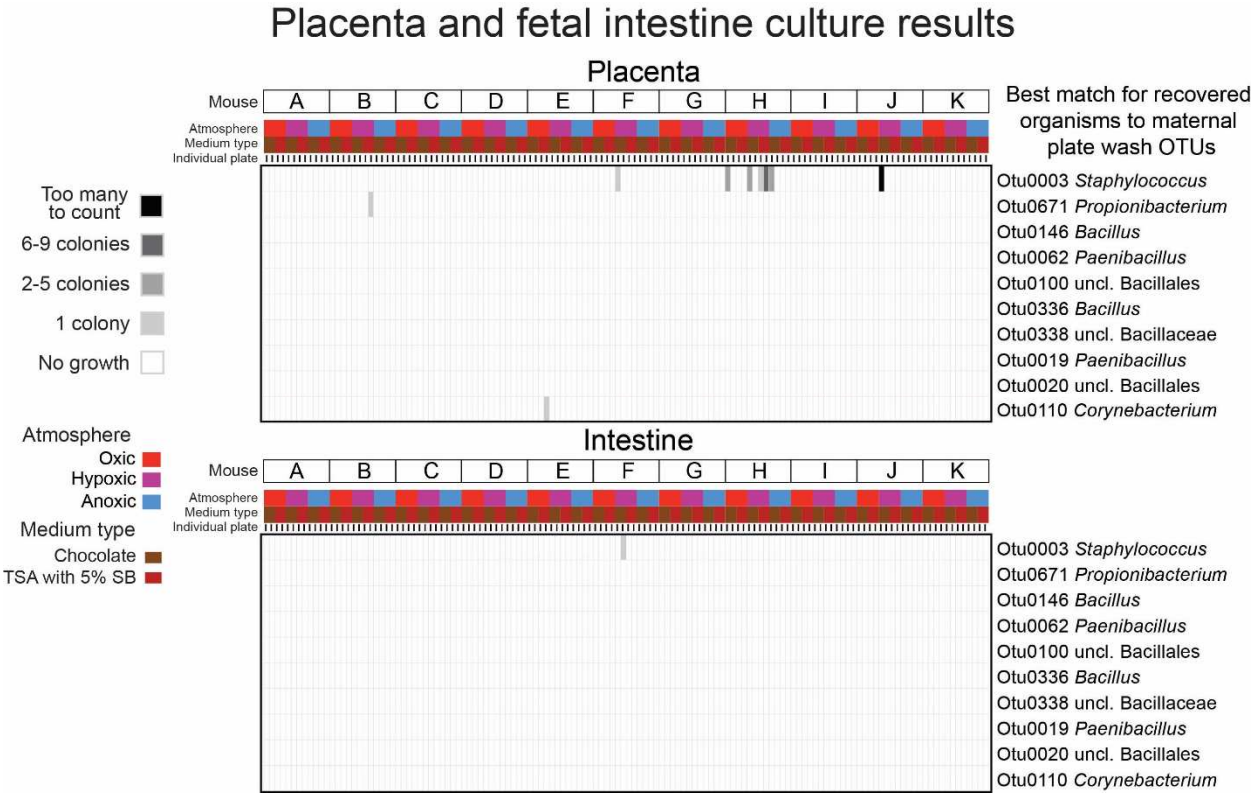

Figure 3.2. Heat maps illustrating bacterial cultivation results for A) placenta and B) fetal intestinal tissues. Each column of the heat map represents a single agar plate. The x-axis indicates the mouse identity, atmospheric condition, growth medium, and paired replicate for each agar plate. The vast majority of blood and chocolate agar plates did not yield any bacterial growth over seven days for placental (93.2%) and fetal intestinal (99.2%) samples. The operational taxonomic units (OTUs) on the y-axis are those that represent the best 16S rRNA gene sequence matches to bacterial isolates recovered from any placental or fetal sample in this study overall (i.e. the OTUs in bold font in Figure 1B).

Table 3.1. Bacterial cultivation results for placental and fetal brain, lung, liver, and intestinal samples.

		Bacterial culture		16S rRNA gene qPCR	16S rRNA gene sequence match between the isolate and ≥ 1 sequence within a 16S rRNA gene library	
Mouse	Placental or fetal body site	Total # of isolates recovered	Top NCBI BLAST taxonomic designation (≥ 99.5% 16S rRNA gene sequence identity unless otherwise indicated)	Was sample bacterial load > that of blank kit controls?	Library for that specific tissue type in that Mouse	Library for any maternal body site in that Mouse
A	Placenta	0		No		
	Lung	0		No		
	Liver	0		No		
	Intestine	0		Yes		
B	Placenta	1	*Cutibacterium acnes*	No	No	No
	Lung	0		No		
	Liver	1	*Cutibacterium acnes*	No	No	No
	Intestine	0		No		
C	Placenta	0		No		
	Lung	1	*Bacillus simplex / frigoritolerans*	No	No	No
	Liver	0		No		
	Intestine	0		No		
D	Placenta	0		No		
	Lung	0		No		
	Liver	0		No		
	Intestine	0		No		
E	Placenta	1	*Corynebacterium tuberculostearicum* (98.5%)	Yes	No	No
	Brain	1	*Bacillus halosaccharovorans*	Yes	No	No
	Lung	0		No		
	Liver	0		No		
	Intestine	0		No		
F	Placenta	1	*Staphylococcus hominis*	No	No	Yes (blood, lung, skin)
	Brain	7	*Bacillus circulans*; *Bacillus megaterium / flexus*; *Bacillus* spp.; Ornithinibacillus sp. Marseille-P3601; *Paenibacillus* spp.	Yes	No	Yes, for 1/7 isolates (skin)
	Lung	0		No		
	Liver	1	Bacillus sonorensis	No	No	No
	Intestine	1	*Staphylococcus hominis* (99.4%)	No	No	Yes (blood, lung, skin)
G	Placenta	0		No		
	Brain	1	*Staphylococcus hominis*	No	No	Yes (peritoneum, skin)
	Lung	0		No		
	Liver	1	*Staphylococcus hominis*	No	No	Yes (peritoneum, skin)
	Intestine	0		No		
H	Placenta	16	*Staphylococcus hominis*; *Staphylococcus epidermidis / caprae / capitis*	No	No	No
	Brain	3	*Staphylococcus hominis*; *Staphylococcus warneri*; *Staphylococcus epidermidis / caprae / capitis*	No	No	No
	Lung	0		No		
	Liver	1	*Paenibacillus timonensis* (98.0%)	No	No	No
	Intestine	0		Yes		
I	Placenta	0		No		
	Brain	0		No		
	Lung	1	*Cutibacterium acnes* (99.0%)	No	No	No
	Liver	0		No		
	Intestine	0		No		
J	Placenta	TMTC*	*Staphylococcus caprae*	No	Yes	Yes (heart, mouth, intestine)
	Brain	0		No		
	Lung	0		No		
	Liver	0		No		
	Intestine	0		No		
K	Placenta	0		No		
	Brain	0		No		
	Lung	0		No		
	Liver	0		No		
	Intestine	0		No		

sites for Mouse H, which included the maternal skin, blood, heart, mouth, lung, liver, proximal intestine, distal intestine, peritoneum, cervix, and vagina **(Table 3.1)**. The placental sample from Mouse J yielded many colonies of *Staphylococcus caprae* on one chocolate agar plate under hypoxic conditions; yet there were no bacterial colonies on the replicate chocolate agar plate incubated under hypoxic conditions or on any other plate for this sample **(Table 3.1, Figure 3.2)**. An exact match of the 16S rRNA gene of this *Staphylococcus caprae* isolate was identified in the 16S rRNA gene survey of placental tissues from Mouse J, as well as in the 16S rRNA gene surveys of the maternal heart, mouth, and proximal intestine samples for Mouse J. However, the bacterial load of the placental sample from Mouse J, as assessed by 16S rRNA gene qPCR, was not high – it was less than the bacterial load of 14/23 (60.9%) DNA extraction kit controls **(Figure 3.1C)**.

Of the 132 blood and chocolate agar plates on which fetal intestinal tissue homogenates were spread, only one yielded growth – a single bacterial colony of *Staphylococcus hominis* **(Figure 3.2)**. The 16S rRNA gene of this bacterial isolate was not detected in the molecular survey of fetal intestines from this mouse (Mouse F), but it was identified in the 16S rRNA gene surveys of maternal blood, lung, and skin from Mouse F **(Table 3.1)**. This sample had the lowest bacterial load of any fetal intestinal sample in the study, and had a bacterial load less than that of 14/23 (60.9%) DNA extraction kit controls **(Figure 3.3)**.

Bacterial culture from maternal compartments

Bacterial cultures of the maternal intestine, mouth, vagina, and lung often yielded lawns of bacterial growth dominated by unclassified Pasteurellaceae, *Lactobacillus*, and *Staphylococcus* **(Figure 3.1B)**. Body site-specific variation in the structure of cultured bacterial communities from maternal samples was evident **(Figure 3.1B)**. For instance, the vast majority

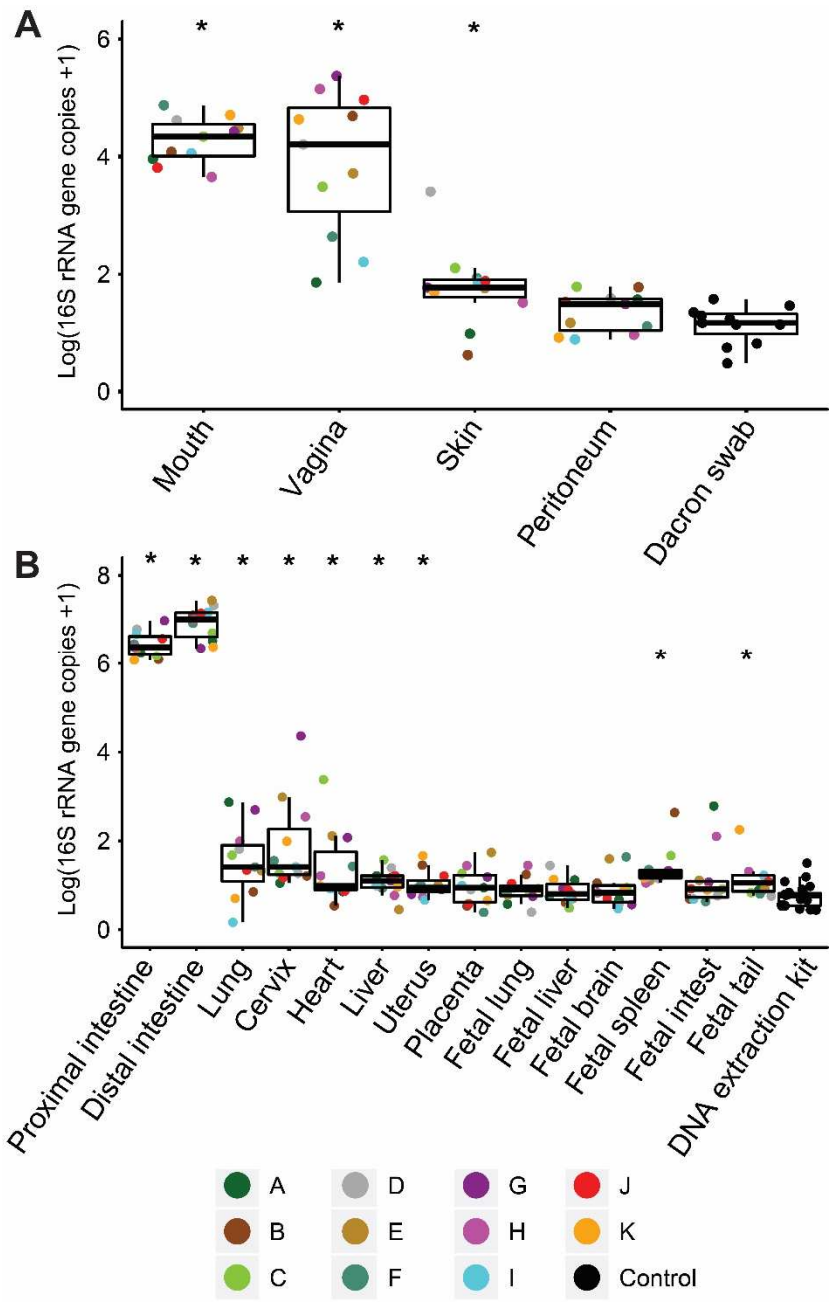

Figure 3.3. Quantitative real-time PCR (qPCR) analyses illustrating variation in bacterial load among A) maternal swab samples and Dacron swab controls, and B) maternal, placental, and fetal tissue samples and blank DNA extraction kit controls. Bars indicate the median and quartile log-16S rRNA gene copy values for each sample and control type. Points, color-coded by mouse identity, indicate the mean values of two replicate qPCR reactions. An asterisk indicates that bacterial loads of that sample type were greater than those of corresponding technical controls.

of bacteria cultured from the vagina were unclassified Pasteurellaceae, while *Bacteroides* and a distinct strain of *Lactobacillus* were consistently cultured from the maternal intestine in addition to the unclassified Pasteurellaceae, *Lactobacillus*, and *Staphylococcus* isolated from other body sites **(Figure 3.1B)**.

Bacterial cultures of the maternal cervix yielded isolates in 6/11 (54.5%) mice **(Table 3.2)**. The most common bacterium cultured from the murine cervix was *Pasteurella caecimuris*; it was recovered in culture from 5/11 cervical samples. In each case, an exact match for the 16S rRNA gene of the *Pasteurella caecimuris* isolate was identified in the 16S rRNA gene survey of the corresponding cervical sample **(Table 3.2)**.

Bacteria were rarely cultured from the uterus (2/11 mice) and maternal liver (4/11 mice) **(Table 3.2)**. The two bacteria cultured from the uterus were *Bacillus niabensis* and *Staphylococcus aureus*. An exact match of the 16S rRNA gene of these isolates was not identified in the 16S rRNA gene surveys of the respective uterine samples. The bacteria cultured from maternal liver samples were primarily *Lactobacillus* and *Staphylococcus* species. Of the nine distinct bacterial morphotypes cultured from maternal liver tissues, only 3 (33%) had an exact match of their 16S rRNA gene identified in the 16S rRNA gene surveys of their respective samples **(Table 3.2)**.

Quantitative real-time PCR (qPCR) of murine and control samples

Bacterial load, as characterized by 16S rRNA gene copy abundance, varied greatly across maternal, placental, and fetal body sites **(Figure 3.3)**. The bacterial loads of swabs of the maternal mouth, vagina, and skin exceeded those of sterile Dacron swabs **(Figure 3.3A)**. Similarly, the bacterial loads of tissues of the maternal proximal and distal intestine, lung, cervix,

Table 3.2. Bacterial cultivation results for maternal cervical, uterine, and liver samples.

		Bacterial culture		16S rRNA gene qPCR	16S rRNA gene sequence match between the isolate and ≥ 1 sequence within a 16S rRNA gene library
Mouse	Low microbial biomass maternal body site	# of unique colony morphotypes recovered	Top NCBI BLAST taxonomic designation (≥ 99.5% 16S rRNA gene sequence identity unless otherwise indicated)	Was sample bacterial load > that of blank kit controls?	Library for that specific tissue type in that Mouse
A	Cervix	1	*Rodentibacter pneumotropicus* (98.0%)	No	Yes
	Uterus	1	*Bacillus niabensis*	No	No
	Liver	0		No	
B	Cervix	0		No	
	Uterus	0		No	
	Liver	2	*Lactobacillus gasseri*; *L. murinus*	No	Yes, for 1/2 morphotypes
C	Cervix	0		No	
	Uterus	0		No	
	Liver	5	*Bacteroides sartorii* (98.0%); *Klebsiella variicola*; *L. gasseri / L. johnsonii*; *L. murinus*; *Staphylococcus hominis*	Yes	Yes, for 2/5 morphotypes
D	Cervix	0		No	
	Uterus	0		No	
	Liver	0		No	
E	Cervix	1	*Pasteurella caecimuris*	Yes	Yes
	Uterus	0		No	
	Liver	0		No	
F	Cervix	0		Yes	
	Uterus	0		No	
	Liver	1	*Staphylococcus epidermidis / S. caprae / S. capitis*	No	No
G	Cervix	5	*Bacteroides sartorii*; *Faecalibaculum rodentium* (97.9%); *Lactobacillus murinus*; *L. reuteri* (99.2%); *Pasteurella caecimuris*	Yes	Yes, for 5/5 morphotypes
	Uterus	0		No	
	Liver	0		No	
H	Cervix	0		Yes	
	Uterus	0		No	
	Liver	0		No	
I	Cervix	6	*Bacillus circulans*; *Pasteurella caecimuris*; *Rodentibacter pneumotropicus* (98.1%); *Staphylococcus hominis*; *S. xylosus*; *Streptococcus thoraltensis* (99.3%)	No	Yes, for 4/6 morphotypes
	Uterus	0		No	
	Liver	0		No	
J	Cervix	1	*Pasteurella caecimuris*	No	Yes
	Uterus	1	*Staphylococcus aureus*	No	No
	Liver	1	*Staphylococcus epidermidis / S. caprae / S. capitis*	No	No
K	Cervix	1	*Pasteurella caecimuris*	Yes	Yes
	Uterus	0		Yes	
	Liver	0		No	

heart, liver, and uterus exceeded those of blank DNA extraction kits **(Figure 3.3B)**. In contrast, bacterial loads of the maternal peritoneum, the placenta, and the fetal lung, liver, brain, and intestine did not exceed those of their respective background technical controls **(Figure 3.3A & B)**. The spleen and tail were the only fetal tissues with bacterial loads exceeding those of blank DNA

extraction kits **(Figure 3.3B)**. However, only 1/11 (9.1%) fetal tail and 2/11 (18.2%) fetal spleen samples had bacterial loads exceeding those of each of the blank DNA extraction kits. Corrected for multiple comparisons, no placental or fetal tissue, including the tail and spleen, had a bacterial load exceeding that of any other placental or fetal tissue (Wilcoxon matched pairs, $p \geq 0.68$).

16S rRNA gene sequencing of murine and control samples

Six of the 23 (26.1%) blank DNA extraction kits, and 8/11 (72.7%) sterile swab controls, yielded a 16S rRNA gene library with ≥ 250 quality-filtered sequences and a Good's coverage ≥ 95%. The prominent (i.e. ≥ 2.25% relative abundance) operational taxonomic units (OTUs) in the bacterial profiles of the DNA extraction kit controls were identified as *Ralstonia*, unclassified Bacillales, *Flavobacterium*, S24-7, *Brevibacterium*, *Pelomonas*, unclassified Bacteroidetes, and *Acinetobacter* **(Figure 3.4)**. However, only two of these prominent OTUs, identified as *Ralstonia* and *Pelomonas*, were present in the bacterial profiles of more than half of the DNA extraction kit controls. A decontam analysis indicated that the OTUs identified as *Ralstonia*, *Pelomonas*, *Pseudomonas*, and *Acinetobacter* were likely background DNA contaminants **(Figure 3.4)**.

The bacterial profiles of placental and fetal samples could not be compared to those of background technical controls because only two of the 77 (2.6%) placental and fetal brain, lung, liver, intestine, spleen, and tail samples included in this study, yielded a 16S rRNA gene library

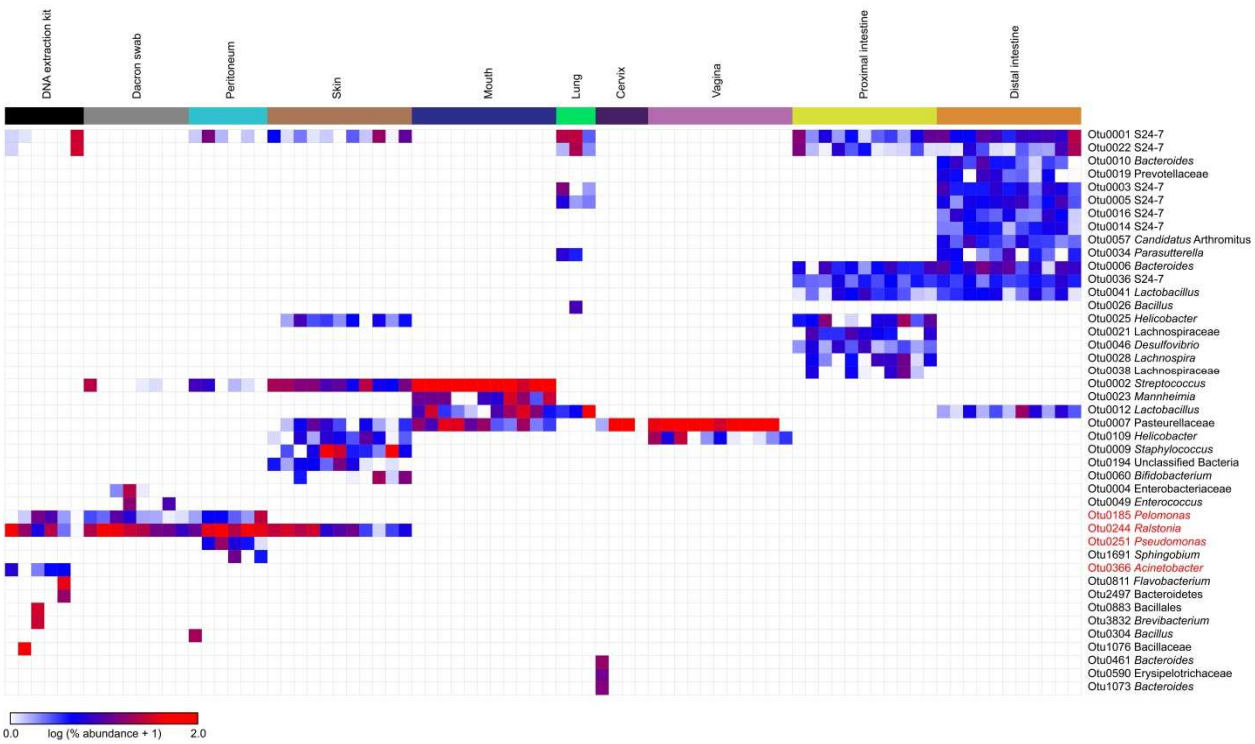

Figure 3.4 Heatmap illustrating the 16S rRNA gene profiles of maternal swab and tissue samples and background technical controls featuring the relative abundances of prominent (≥ 2.25% average relative abundance) operational taxonomic units (OTUs). The four OTUs in red font were identified as background DNA contaminants by the R package decontam.

with ≥ 250 sequences and a Good's coverage ≥ 95%. These two samples were the placenta from Mouse I and the fetal spleen from Mouse B. The placenta from Mouse I had an average bacterial load in comparison to that of other placentas **(Figure 3.3)**, and no bacteria were cultured from the placental tissues of this mouse **(Table 3.1, Figure 3.1, Figure 3.2)**. The prominent OTUs in the bacterial profile of the placental sample from Mouse I were identified as *Bacteroides*, *Akkermansia*, S24-7, *Lactobacillus*, and *Escherichia*. The fetal spleen from Mouse B had the highest bacterial load of any fetal spleen sample; its bacterial load was 58% higher than any other spleen sample **(Figure 3.3)**. The prominent OTUs in the bacterial profile of the fetal spleen from Mouse B were *Lactobacillus*, S24-7, and unclassified Lachnospiraceae.

All maternal skin, mouth, proximal and distal intestinal samples yielded a 16S rRNA gene library with ≥ 250 sequences and a Good's coverage ≥ 95%. Six (54.5%), four (36.4%), and three (27.3%) maternal peritoneal, cervical, and lung samples, respectively, yielded a 16S rRNA gene library with ≥ 250 sequences and a Good's coverage ≥ 95%. However, no maternal liver or uterine samples, and only one (9.1%) maternal heart sample, yielded a 16S rRNA gene library. The structure of the bacterial profiles of the maternal body sites with at least three 16S rRNA gene libraries were compared with those of background technical controls **(Figure 3.4, Figure 3.5)**.

The taxonomic identities of prominent OTUs varied among maternal body sites **(Figure 4)**. Maternal proximal and distal intestinal samples had the most OTU-rich bacterial profiles. The maternal proximal intestine was characterized by *Bacteroides*, *Desulfovibrio*, *Helicobacter*, *Lachnospira*, unclassified Lachnospiraceae, *Lactobacillus*, and S24-7, while the maternal distal intestine had bacterial profiles consistently comprised of "*Candidatus* Arthromitus," *Bacteroides*, *Lactobacillus*, *Parasutterella*, unclassified Prevotellaceae, and S24-7. Maternal

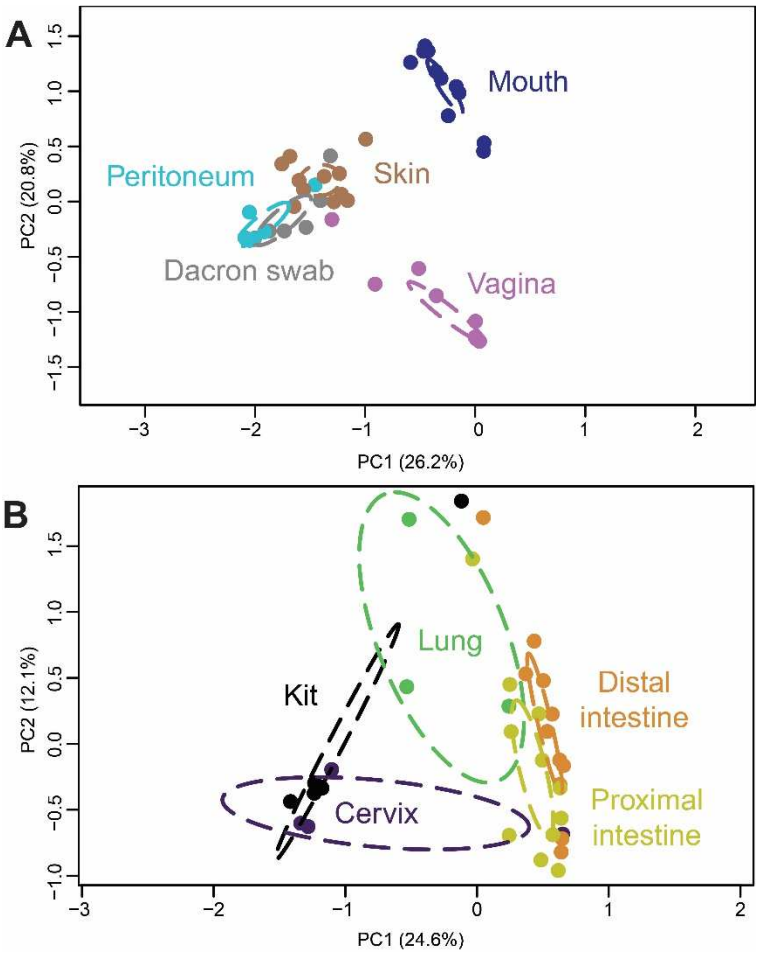

Figure 3.5. Principal Coordinates Analysis (PCoA) of maternal samples and controls illustrating variation in 16S rRNA gene profiles among A) maternal swab samples and Dacron swab controls, and B) maternal tissue samples and blank DNA extraction kit controls. 16S rRNA gene profiles were characterized using the Bray-Curtis similarity index.

vaginal and cervical bacterial profiles were dominated by unclassified Pasteurellaceae; the vagina also consistently contained *Helicobacter*. Maternal lung bacterial profiles were typified by *Lactobacillus* and S24-7, while those of the maternal mouth were dominated by *Streptococcus*, *Mannheimia*, *Lactobacillus*, and unclassified Pasteurellaceae. Maternal skin, a low microbial biomass site **(Figure 3.3A)**, and the peritoneum, a very low to nonexistent microbial biomass site **(Figure 3.3A)**, had bacterial profiles that overlapped with those of background technical controls more so than did the profiles of higher microbial biomass sites **(Figure 3.4A)**. Specifically, skin bacterial profiles consistently contained *Bifidobacterium*, *Helicobacter*, unclassified Pasteurellaceae, *Ralstonia*, S24-7, *Staphylococcus*, and *Streptococcus*. *Ralstonia* was the dominant OTU in the bacterial profiles of the maternal peritoneum, as well as in the profiles of the background technical controls **(Figure 3.4)**. Indeed, the bacterial profiles of the maternal peritoneum were not distinguishable from those of background technical controls (Bray-Curtis similarity index; NPMANOVA, $F = 0.974$, $p = 0.467$) **(Figure 3.5)**.

Comprehensive consideration of individual placental and fetal tissues across microbiological inquiries

Of the 165 total bacterial cultures of placentas from the 11 mice, only nine (5.5%) yielded even a single colony, and five of those nine positive cultures came from a single mouse; 2) of the 165 total bacterial cultures of fetal intestinal tissues, only one (0.6%) was positive, yielding a single isolate of *Staphylococcus hominis*; 3) the bacterial loads of placental and fetal brain, lung, liver, and intestinal samples were not higher than those of DNA extraction kit controls; 4) only two (2.6%) placental or fetal tissue samples yielded a 16S rRNA gene library with at least 250 sequences and a Good's coverage value of 95%; 5) the 16S rRNA gene libraries of each maternal skin, mouth, vaginal, and proximal and distal intestinal sample met these criteria, as did at least 25% of maternal lung, cervical, and peritoneum samples; 6) similar to the

placental and fetal tissues samples, maternal heart, liver, and uterine samples did not yield 16S rRNA gene libraries with at least 250 sequences and a Good's coverage value of 95%; 7) overall, from all placental or fetal tissues, there was only a single bacterial isolate (*Bacillus circulans*, cultured from the fetal brain tissue of Mouse F) that was cultured from a placental or fetal tissue that had a bacterial load higher than that of background technical controls, and that was identified in the 16S rRNA gene surveys of at least one of that fetus' maternal samples **(Table 3.1)**.

V. Discussion

As the era of microbiome discovery and research progresses out of its infancy, many gaps of knowledge still exist. The Human Microbiome Project [42] was the first concerted and collaborative effort to describe the healthy microbiota of humans, and while many environments have had their microbiota characterized, many animals have not received the same degree of attention in terms of their microbiota. While there are many reports of various microbial communities for a wide range of animals, they are rarely described to the degree seen in the HMP, and even rarer, for multiple body sites of a particular animal to be described. Various issues, such as a lack of standardized methodologies and tools for describing microbiota, are difficult to overcome given the current limitations of technology and the relatively early stage of the field. One seminal effort was made to describe a portion of the murine microbiota, that of the intestine [224]. Investigators utilized culture, metagenomics, and 16S rRNA gene sequencing to extensively characterize bacteria associated with the mouse intestine. Through culture, they identified a new family, ten genera, and four species, as characterized by 16S rRNA gene sequencing, and even found two species with high prevalence among multiple strains of mice from several animal facilities. The importance of a fully characterized and understood mouse

microbiota is undeniable given that the majority of animal-driven, human-translatable research relies on the mouse model.

Despite this limited knowledge of the broader murine microbiome, several studies have focused on investigating the murine placenta and fetal intestine for evidence of a microbiota. For instance, Martinez et al. [214] performed bacterial culture, 16S rRNA gene qPCR, and 16S rRNA gene sequencing on the placenta and fetal intestines of 13 mice at day 17 of gestation (same gestational age was used in our study). No bacteria were recovered from culture surveys; however, they did find higher bacterial loads of the fetal intestine to be higher than those of the placentas via qPCR. The bacterial profiles of placenta and fetal intestines were found to be different from each other after removing any OTUs detected in the background technical control samples of the 16S rRNA gene surveys. While Martinez et al. found no evidence that murine fetuses are populated by microbial communities, they are exposed to bacterial DNA *in utero*. Although limited to molecular surveys, another study by Kuperman et al. [225] investigated 24 murine placental samples at gestational day 19 and was unable to detect 16S rRNA gene amplicons after PCR, similar to the limited detection of 16S rRNA gene signals we found in our study [212].

In contrast, Younge et al. observed bacterial signals in the murine placenta and fetal intestines from mice at early, mid, and late gestations. Bacteria cultured predominantly from fetuses at mid-gestation and were most commonly *Lactobacillus, Escherichia, Enterococcus, Bacteroides,* and *Bacillus*. Additionally, bacteria were visualized in fetal intestines using fluorescent *in situ* hybridization (FISH) with a universal probe targeting the bacterial 16S rRNA gene. The possibility that the cultured and visualized bacteria originated from maternal tissues due to contamination during sampling is unlikely and was addressed in mechanistic studies

utilizing surgical techniques and inoculations with two different antibiotic-resistant strains of *E. coli*. Using 16S rRNA gene sequencing, the bacterial profiles of placenta and fetal intestines were similar to each other, while variation was observed when comparing early gestation samples to mid and late gestation samples. Sourcetracker analyses indicated that the bacterial signals from early gestation fetal intestines were most likely originating from background technical controls or unknown sources, while mid and late gestation samples were attributed to the placenta or amniotic membrane. Younge et al. [213] concluded that fetal exposure to microbial communities is occurring, and that the exposure is likely coming from the placenta and extraplacental membranes *in utero*.

Similar to the debate over a microbiota in the human placenta, there is contradictory evidence regarding a murine placental microbiota and *in utero* colonization. The data presented in this chapter are more consistent with the prior reports of Martinez et al. [214] and Kuperman et al. [225] and are contradictory to the evidence reported by Younge et al. [213] regarding *in utero* colonization. In our study, culture of bacteria from placental and fetal tissues was generally rare. Most of the bacterial isolates were identified as *Staphylococcus hominis*. The origin of these bacteria could be maternal sites, as *Staphylococcus* spp. were routinely cultured from maternal sites and *Staphylococcus hominis* specifically was identified in molecular surveys of the maternal skin. Alternatively, these bacteria could potentially be contaminants from laboratory personnel, given that two of the five bacterial isolates recovered from negative control plates in this study were also *Staphylococcus hominis*. The other bacteria (*Bacillus*, *Corynebacterium*, *Paenibacillus*, and *Propionibacterium*) cultured from placental and fetal samples were rarely, if ever, cultured from maternal samples or identified in the molecular surveys of maternal samples. Given that the only possible source of placental and fetal microbiota is microorganisms in the

maternal compartments, the latter finding suggests that these bacteria were likely contaminants. Furthermore, there was no consistent recovery in culture of specific microorganisms (aside from *Staphylococcus hominis*) from multiple placental and fetal tissues from the same fetus or in the same tissue types among fetuses from different litters. Notably, the taxonomic identities of bacteria cultured in the current study generally differed, with the exception of *Staphylococcus*, *Bacillus*, and *Paenibacillus*, from those initially reported by Younge et al. [213] in placental and fetal tissues. Therefore, across studies culture has not provided consistent evidence for a placental or fetal microbiota.

Furthermore, qPCR revealed that the bacterial loads of the placenta, fetal lung, liver, brain, and intestine did not exceed those of background technical controls, whereas samples from maternal sites, excluding the peritoneum, did exceed those of controls. In addition, there was no variation in bacterial load among placental and fetal tissues. These results are in contrast to those of Martinez et al. [214] in which the bacterial loads of the fetal intestine exceeded those of the placenta. To our knowledge, no other studies have directly compared the bacterial loads of the placenta and fetal intestine in mammals. However, the qPCR results in our study agree with prior qPCR investigations of human placental tissues – the bacterial loads of placentas are indistinguishable from those of background technical controls [128, 150, 167]. Hence, there remains disagreement among studies with respect to the extent of bacterial biomass in placental and fetal tissues.

Herein, the murine placenta and fetal tissues did not yield substantive 16S rRNA gene sequence libraries, while the maternal sites other than the uterus, heart, and liver consistently did so. These results are consistent with those of Kuperman et al. [225], in which 30 cycles of PCR did not yield discernible amplicons from murine placental tissue. Notably, in our study, triple

library preparations were performed and pooled for each sample, and still minimal amplicons were generated after 30 cycles of PCR. Martinez et al. [214] also used 30 cycles of PCR in their sequence library preparations and included samples in their analyses if they yielded at least 200 quality-filtered sequences, reporting a distinct bacterial DNA signal in the placenta and fetal intestine. In this study, we only included samples in analyses if they yielded at least 250 quality-filtered sequences with a Good's coverage value of at least 95%. If we had used the criterion of 200 sequences, independent of any consideration of Good's coverage, only one additional fetal sample would have been included in analyses. Younge et al. [213] generated substantive sequence libraries for placental and fetal intestine samples; however, their library preparation protocol was based on that of the Earth Microbiome Project (i.e. 35 cycles of PCR). The discrepancies among murine studies may therefore be due to underlying differences in the sequence library protocols used. Nevertheless, as with culture and qPCR approaches, we did not find consistent evidence of a bacterial signal in placental and fetal tissues using DNA sequencing.

Notably, in this study, there was only one case in which a bacterial isolate (i.e. *Bacillus circulans*) from a placental or fetal sample (i.e. fetal brain) had a bacterial load exceeding that of all background technical controls, and in which the bacterium was also identified in molecular surveys of at least one corresponding maternal sample (i.e. maternal skin). Therefore, in this one case, there may have been hematogenous transfer from a distant maternal site to the fetus. However, overall, there was not consistent evidence of resident bacterial communities in the murine placenta or the fetus.

Although the context of this chapter has shifted from humans to mice, it is important to highlight that our work and the works of others [213, 214, 225] are challenging paradigms of

sterility. That being stated, the value of finding consistent and reproducible evidence cannot be neglected, nor can the caveats of working with low microbial biomass tissues. Our study emphasizes the importance of including appropriate background technical controls, as well as positive and negative tissue controls, in all microbiological approaches from culture to sequencing when reevaluating paradigms of sterility. Ultimately, while several studies, including our own, have included multiple methodologies of microbiologic inquiry, no studies have thus far met all the criteria put forth in Theis et al. [150] for establishing the presence of a resident microbiota. One major oversight in the work thus far has been the lack of inclusion of a germ-free mice cohort compared to a wild type cohort, which presents an opportunity to investigate *in utero* exposure and colonization that is uniquely available to animal studies and should be included in future work.

CHAPTER 4: A URINARY MICROBIOTA IN PREGNANCY: CULTIVATION- AND MOLECULAR-BASED COMPARISON OF FOLEY CATHETERIZED URINE, CLEAN CATCH URINE, AND VAGINAL SWABS FROM THE SAME WOMEN

Note: Ali Alhousseini, a former PhD student in the Department of Physiology and the Theis laboratory at Wayne State University contributed equally to this chapter. I am responsible for all analyses, generation of tables and figures, writing of the Abstract, Methods, and Results, and revision of the Introduction and Discussion. Dr. Alhousseini collected all the clinical samples, participated in the processing of these samples, and originally wrote the Introduction and Discussion sections. A prior version of this study was included in his dissertation thesis as well.

I. Abstract

Urine and the bladder have historically been considered sterile, especially in the context of clinical assessment. Recent work has been capitalizing on enhanced culture techniques and next-generation molecular sequencing surveys to re-assess the sterility of urine. Indeed, the scientific community has been shifting towards accepting that the bladder does contain microorganisms in healthy individuals, yet further investigation is warranted, as demonstrated in Chapters 2 and 3 of this dissertation. My focus these past five years has been on paradigms of sterility in perinatal medicine, and studies on the urine of pregnant women have been limited thus far. This is surprising given that pregnant women are at an increased risk of urinary tract infections (UTIs) and UTIs can lead to pregnancy complications. Because of these vulnerabilities, investigating the bladders and urine of pregnant women is critical in evaluating shifting paradigms of urine sterility.

To better understand the existence and potential role of a bladder microbiota during pregnancy, we must first establish suitable approaches to its study and characterize which bacteria inhabit the bladder. To do so, we characterized the urinary microbiota of 25 pregnant women (delivering after 35 weeks gestation) by comparing the bacterial profiles of paired catheter urine, clean catch urine, and vaginal swabs using cultivation and molecular

microbiological survey methods. For culture, three bacterial taxa were detected in at least 20% of all urine samples (*Lactobacillus* species, coagulase negative *Staphylococcus* species, and *Ureaplasma urealyticum*), and all three taxa were detected less frequently in Foley catheter urine than in CC urine. *Ureaplasma urealyticum* was the most frequently recovered bacteria in Foley catheter urine (13/25 women). 16S rRNA gene surveys showed that the microbial profiles of Foley catheter urine and vaginal swabs differed in composition and structure, but that the profiles of clean catch urine and vaginal swabs were similar. For all three sample types, bacterial profiles were abundant in *Lactobacillus* and *Gardnerella* species, but there was variation in lower abundance taxa among these three sample types, especially with regards to Foley catheter urine, which is presumably most closely representative of a bladder microbial community, if one exists.

Overall, our data suggest that residential bacterial communities exist in the female bladder and urine during pregnancy, and that there is overlap between those communities and those in the vagina. Aside from *Lactobacillus* and *Gardnerella* species, which were frequently detected in molecular surveys of Foley catheter urine obtained directly from the bladder, there was high inter-individual variability of less abundant taxa. Several lower abundance taxa that were differentially more abundant in catheter urine than vaginal swabs in molecular microbiology surveys were *Ureaplasma*, *Finegoldia, Anaerococcus,* and *Fenollaria* species. Of these, *Ureaplasma* was detected by cultivation in a majority of women. The remaining bacteria are anaerobes, which would require targeted or enhanced culture methods for their detection. Future directions include validating these findings in a larger cohort and pursuing more robust cultivation methods that are efficient for capturing low abundance anaerobes in clinical samples, especially Gram-positive anaerobic cocci (including *Finegoldia, Anaerococcus, Peptoniphilus,* and *Atopobium)*. Additionally, investigation of the bladder of women delivering preterm

(especially early preterm, < 32 weeks of gestation) is warranted for comparisons of bladder microbial communities to women delivering at term will allow us to assess influences of resident bacterial communities on perinatal health and pregnancy outcomes, which can ultimately be used to identify specific bacteria or bacterial communities whose presence or absence can serve as potential risk indicators for adverse pregnancy outcomes, especially spontaneous preterm birth.

II. Introduction

The bladder and urine have historically been regarded as sterile [115, 169]. However, this perception is being reconsidered [69, 114, 117, 118, 133, 226, 227]. Given appropriate cultivation conditions, based on atmospheric and metabolic requirements, microorganisms can be cultured from urine in healthy patients [64, 228]. Furthermore, capitalizing on contemporary advances in next-generation sequencing technologies, urinary microorganisms have been identified and characterized among asymptomatic non-pregnant women [69, 114, 117, 118, 133, 226, 227], and, in a single study, among pregnant women [169]. Collectively, these studies have suggested that bacteria reside in the human bladder as commensals, and even potentially as mutualists [69]. For example, a current hypothesis is that microbes residing within the bladder and urine of healthy people competitively exclude potential pathogens and that dysbiosis of these resident microbial communities could lead to an overgrowth of opportunistic pathogens, resulting in urinary tract infections (UTIs), urinary urge incontinence, and other urinary tract disorders [63, 64, 67].

Despite recent work, the existence of a urinary microbiota has not yet been effectively investigated using contemporary molecular microbiology techniques in the context of pregnancy outcomes. Bacterial presence in the urinary tract has been considered pathologic and has been classified into asymptomatic bacteriuria (ASB), urinary tract infection, or pyelonephritis [137,

169, 229, 230]. Pregnant women are an important population in which to assess any potential influences of a urinary microbiota due to vulnerabilities associated with the urinary tract during pregnancy [162, 163, 231]. Specifically, women experience physiological and morphological changes during pregnancy, including ureteral dilation, decreased bladder tone, displacement and compression of the bladder, increased renal length, and decreased peristalsis. The consequences of these alterations can include urinary stasis, vesicoureteral reflux, and hydronephrosis [162, 231], each of which contribute to conditions that are conducive to microbial growth [163], thereby increasing the risk of ascending UTIs. In fact, UTIs are the most common bacterial infection in women during pregnancy, occurring in up to 8% of pregnancies, with approximately 5% of women experiencing at least one UTI event during a given pregnancy [232]. Urinary tract infections can lead to significant maternal and perinatal complications, including preterm birth, low birth weight, maternal sepsis, and disturbance of the immune system, including an increase in inflammation [157, 158, 233-243]. While women diagnosed with having asymptomatic bacteriuria (ASB) are at similar risk for adverse pregnancy outcomes as UTIs, the main difference between these two diagnoses, aside from symptom presentation, has been the magnitude of detectable bacteria in urine, or colony forming units (CFUs) from urine culture, which is still the clinical diagnostic standard along with urinalysis. Notably, asymptomatic bacteriuria has been reported in up to 10% of pregnancies [244-248], and, if left untreated, ASB can lead to symptomatic UTIs, including pyelonephritis in 30-40 % of the cases [249]. A recent Cochrane review therefore recommended treatment of ASB to reduce the incidence of pyelonephritis during pregnancy [244]. This is in line with typical clinical practice as a positive urine culture (> 10,000 colony forming units per ml) during pregnancy has traditionally elicited

antibiotic treatment and a repeat of culture within 1-2 months to confirm resolution of the infection [137].

An important consideration that needs to be addressed is that asymptomatic bacteriuria, and to a lesser extent UTIs, have historically hinged on the long-held belief that the upper urinary tract is sterile. While this now stands contradictory to multiple studies that have been published using enhanced culture and 16S rRNA gene sequencing surveys suggesting the existence of a urinary tract or bladder microbial community in healthy individuals, there are two important caveats inherent in characterizing urinary microbiota profiles. First, urine samples are susceptible to vulvovaginal contamination, so there is risk that characterized microbes were not actually residing in the bladder or the urine. Studies by Wolfe and colleagues [48, 115, 118] found many genera, although the most represented were *Lactobacillus, Prevotella, Staphylococcus, Atopobium, Corynebacterium, and Gardnerella*, which are all commonly associated with the vagina and human skin. Second, if there are indeed resident urinary microbial communities, they are present in very low abundances and thus, when characterizing them through next-generation sequencing, there is risk of amplifying and characterizing background bacterial DNA contamination from DNA extraction kits and PCR reagents [98, 99, 128, 149, 189]. 16S rRNA gene sequences are ubiquitous in the environment and do not indicate viability of any surveyed bacteria. From our analysis of the current literature, we found that most of the publications investigating the urinary tract microbiota did not include appropriate technical controls to properly account for background DNA. Many of the less abundant genera detected in recent studies represent species that are not captured by routine cultures or have never been successfully cultured. It is also important to perform viability assays, such as expanded quantitative urine culture (EQUC) methods, to demonstrate viability of bacteria from these

samples because molecular surveys do not differentiate between ubiquitous environmental 16S rRNA gene sequences and those from living bacteria [115]. Additionally, some of the data provided in studies were low in sample size (e.g., n = 2), and conducted viability assays on only a subset of samples [117, 118]. Therefore, it is important for future work in this area to include larger sample sizes, more rigorous and extensive culture techniques, and ample background technical controls to draw a proper conclusion on the existence of microbial communities in the bladder.

Our broad objective is to further investigate and resolve the ambiguities surrounding the presumed presence of urinary tract microbial communities among normal asymptomatic pregnant women, and to ultimately ascertain potential influences of these communities on women's reproductive health. By investigating the presence of microbial communities in the bladder of pregnant women, follow-up studies can elucidate any associations that certain microbial species or community structures may have with pregnancy outcomes. Should we find that urine samples are not significantly different from technical controls, the medical community can reassess diagnostic criteria to address the asymptomatic presence of bacteria as an indicator of subclinical infection that can be addressed with prophylactic measures. Additionally, this result should reduce or stop the publication of studies alluding to urinary microbiota without including necessary technical controls. Conversely, if we find that there is a resident microbial community within the female urinary tract, it is critical that current paradigms of a sterile urinary tract be revised to accommodate these resident microbes and the corresponding microbiome. Additionally, given that infection in non-sterile sites may indicate disruption of microbial homeostasis, understanding of the underlying microbial community is critical for addressing and preventing urinary tract infections. This would also render the term bacteriuria obsolete and

require new approaches for handling the updated context of urinary tract disorders. In particular, women's reproductive health serves to benefit most by this paradigm shift through better risk indicators through screening of these communities. If we find distinct community types associated with particular disease states or increased risks for obstetric or gynecologic complications, preventative measures can be taken to either alter the community or replace it with probiotics, antibiotics, or a combination of the two (i.e. synbiotics).

The specific and principal objective of this study was to determine if there is a viable microbiota in the bladder and urine during pregnancy and, if so, to characterize it. Secondary objectives were to: 1) assess similarity of culture and next-generation sequencing characterizations of urinary microbial profiles; 2) assess the microbial load of urine sampled using catheter and clean catch collection methods via quantitative real-time PCR; 3) compare the composition and structure of the microbial profiles of urine from pregnant women obtained using catheter and clean catch collection approaches with those of background technical controls; and 4) contrast the microbial profiles of the urine of pregnant women obtained through these two collection approaches with those of the vagina to assess potential vulvovaginal contamination. These objectives were achieved by collecting 3 sample types, catheter urine, clean catch urine, and vaginal swabs, from pregnant women and characterizing their urine and vaginal microbiota profiles through both culture and next-generation sequencing surveys. To accomplish our objectives, our study was divided into two components. In Study Component 1, we evaluate whether the urine of pregnant women has bacterial loads and bacterial profiles distinct from those of technical controls, as assessed by quantitative real-time PCR and 16S rRNA gene sequencing. Additionally, we determine what is an appropriate volume of urine to use for effective surveys of potential bladder microbial communities. In Study Component 2, after

establishing that urine samples are distinct from technical controls and identifying an appropriate urine volume for assessing microbial communities, we compare urine collection methods to assess the most informative and accurate method for detecting and characterizing microorganisms in the bladder, while controlling for vulvovaginal and background DNA contamination. Finally, we detail the evidence of a bladder microbiota in this cohort and identify the bacteria that are likely members of the bladder microbiota during pregnancy.

III. Methods and Materials

Clinical specimens

Urine and vaginal swab samples were obtained at the Perinatology Research Branch, an intramural program of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, U.S. Department of Health and Human Services, Wayne State University (Detroit, MI), and the Detroit Medical Center (Detroit, MI). The collection and use of human materials for research purposes were approved by the Institutional Review Boards of the National Institute of Child Health and Human Development and Wayne State University. All participating women provided written informed consent prior to sample collection.

Study design

This was a cross-sectional study in which the urinary and vaginal microbiota were examined in 25 women admitted for delivery after 35 weeks gestation. There were two components to the study. First, the bacterial loads and bacterial profiles of urine samples collected using Foley catheter and clean catch sampling methods were compared in a subset of women (n = 8). These comparisons were made across a range of urine volumes (1 ml, 1.8 ml,

5.4 ml, 10 ml, and 25 ml).Second, the bacterial loads and bacterial profiles of 5.4 ml of Foley catheter urine, clean catch urine, and vaginal swabs were compared across all women and were contrasted with those of background contamination controls.

Inclusion and exclusion criteria

Inclusion criteria: 1) delivery after 35 weeks of gestation, and 2) intact membranes at the time of collection of vaginal swabs and clean catch urine samples. Exclusion criteria: 1) any maternal or fetal condition that requires termination of pregnancy; 2) known major fetal anomaly or fetal demise; 3) active vaginal bleeding; 4) serious medical illness (e.g. renal insufficiency, congestive heart disease, chronic respiratory insufficiency, etc.); 5) asthma requiring systemic steroids; 6) patient requiring anti-platelet or non-steroidal anti-inflammatory drugs; 7) active hepatitis; and 8) signs or symptoms of urinary tract infection, asymptomatic bacteriuria, and pyelonephritis at the time of sampling.

A urinary tract infection (UTI) is bacterial growth of more than 10,000 colonies of a single bacterial type per milliliter (CFU/ml) of urine coincident with one of the following symptoms: hematuria, dysuria frequency, urgency or suprapubic pressure [229, 230]. However, because others have recommended a much lower threshold of 100 CFU/ml for diagnosing symptomatic UTIs [250], we applied the latter definition. Asymptomatic bacteriuria (ASB) is the presence of 100,000 CFU/ml without any associated symptoms [137, 229, 230]. Pyelonephritis is infection of the kidneys and the presence of systemic signs or symptoms such as fever, nausea and vomiting, chills or flank pain [137, 229, 230]. Again, no woman in this study had a urinary tract infection, asymptomatic bacteriuria, or pyelonephritis.

In the first component of the study, which assessed bacterial load and profiles of urine samples at multiple volumes compared to background contamination, no subject had received

antibiotics in the last week. In the second component of the study, evaluating differences in bacterial load and profiles of Foley catheter urine and clean catch urine and vaginal swabs, no subject had received antibiotics in the last month.

Sample collection

On admission, each woman provided a mid-stream clean catch urine specimen (CC). A speculum exam was performed, and a sample of vaginal fluid was collected from the posterior vaginal fornix under direct visualization by an obstetrician using a FLOQSwab (Copan Diagnostics, Murrieta, CA, USA). During labor or prior to a cesarean delivery, a sterile Foley catheter was inserted, and a second urine specimen was collected (Foley catheter was placed 8.36 ± 1.93 (mean ± SE) hours after the clean catch sample was collected). Urine (excluding the aliquot for culture, see below) and vaginal swabs were frozen at -80°C within one hour of collection.

Bacterial culture of urine

A 2 ml aliquot of urine was sent for bacterial culture. Aliquots of urine were delivered to the University Laboratories Microbiology Core in the Detroit Medical Center, wherein they were processed and cultured under aerobic and anaerobic conditions that day. A genital mycoplasma assay (Mycofast US; Logan, UT) was also conducted for each urine sample [251]. Incubation for aerobic, anaerobic and *Mycoplasma*/*Ureaplasma* cultures was performed at 35°C. Aerobic plates were TSA 5% SB (Trypticase Soy Agar w/5% Sheep's Blood), Columbia CNA SB, MacConkey and MTM II (Modified Thayer Martin). Anaerobic plates used were Brucella OxyPRAS Plus, KVL/BBE Biplate (Brucella Laked Blood Agar with Kanamycin and Vancomycin/Bacteroides Bile Esculin Agar) and CDC ANA BLD (CDC Anaerobic Blood Agar). Aerobic cultures were grown in an incubator with 8% CO_2, anaerobic cultures were grown in a plastic anaerobic culture

chamber in a non-CO_2 incubator, and the *Mycoplasma/Ureaplasma* cultures were grown in an oxic environment without CO_2 supplementation. In each case, one drop of urine, equivalent to approximately 0.05 ml, was used. Urine samples were incubated for 48 hours. The taxonomies of resultant isolates were characterized using Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) within the University Laboratories Microbiology Core [252].

Genomic DNA extractions

Preparation of urine samples for DNA extraction: For 1, 1.8, and 5.4 ml sample volumes, DNA extractions were performed as follows: Urine samples were originally stored at -80°C in either 2 ml cryovials or in 15 ml centrifuge tubes. Samples were thawed at room temperature and thoroughly vortexed before aliquoting into 1.8 ml mini-centrifuge tubes (one tube for 1 and 1.8 ml sample volumes, and three tubes for the 5.4 ml sample volume). Samples were spun in a mini-centrifuge in a 4ºC cold room for 30 minutes at 17,000 *g*. After centrifugation, each sample had the majority of supernatant removed. For the 1 ml sample, approximately 750 µl of supernatant was carefully removed with a 1 ml pipette tip, avoiding the pellet, thereby leaving about 250 µl of the supernatant and the pellet for DNA extraction. For the 1.8 ml sample, 775 µl was removed twice, again being careful to avoid disturbing the pellet, leaving about 250 µl of the supernatant and the pellet for DNA extraction. For the three 1.8 ml tubes constituting the 5.4 ml sample, 860 µl was removed twice from each tube, carefully avoiding the pellet, leaving 80 µl of supernatant and the pellet in each tube for DNA extraction. The initial step of the DNA extraction protocol requires adding 500 µl of the kit's PowerBead Solution to the sample; the PowerBead Solution was added directly to these 1.8 ml tubes. The tubes were then thoroughly mixed through vortexing and by pipetting the solution up and down to ensure that the pellet was

dislodged into solution and would be transferred to the bead tube in the next extraction step. For the 5.4 ml sample, 500 µl of PowerBead Solution was added to the first 1.8 ml tube, the tube was then mixed, transferred to the second tube, mixed, transferred to the third tube, and mixed again before being transferred to the PowerBead Tube in the next extraction step.

For DNA extractions performed on 10 and 25 ml samples: Urine samples were originally stored at -80°C in either 2 ml cryovials or in 15 ml centrifuge tubes. Samples were thawed at room temperature and thoroughly vortexed before transferring 10 or 25 ml into 50 ml centrifuge tubes. These samples were spun at 4ºC at 17,000*g* for 30 minutes. After centrifugation, supernatant was removed without disturbing the pellet. The initial step of the DNA extraction protocol requires adding 500 µl of PowerBead Solution to the sample, so the PowerBead Solution was added directly to these 50mL tubes. These tubes were then thoroughly mixed through vortexing and pipetting the sample up and down to ensure that the pellet was dislodged into solution and would be transferred to the bead tube for the following step in the extraction protocol.

Extraction protocol: Genomic DNA was extracted from urine and vaginal swab samples using QIAGEN DNeasy PowerLyzer PowerSoil Kits according to the manufacturer's protocol with the following modifications: 1) instead of adding 750 µl of PowerBead Solution to each sample, 500 µl of PowerBead Solution and 200 µl of phenol/chloroform:isoamyl alcohol were added and the sample was incubated in the PowerBead Tubes at room temperature for 10 minutes, 2) steps that entail adding Solutions C2 and C3 were combined into one step; 1 µl of RNase A enzyme was also added, 3) instead of adding 1200 µl of Solution C4, 650 µl of C4 and 650 µl of 100% ethanol were added, 4) the dry-spin after Solution C5 was extended from 1 to 2 minutes, 5) Solution C6 was heated to 60°C prior to elution of DNA, and 6) 60 µl instead of 100 µl of

Solution C6 were added to the Spin Column and incubated for 5 minutes before final centrifugation. Blank DNA extraction kits with no urine sample added (n = 12) were processed alongside urine samples. Purified DNA was stored at -20°C.

Quantitative real-time PCR (qPCR) of 16S rRNA genes in samples

Bacterial DNA abundance within samples was determined via quantitative real-time PCR (qPCR) amplification of the V1 – V2 region of the 16S rRNA gene according to a protocol described by Dickson et al [70] with minor modifications. These modifications included the use of a degenerative forward primer (27f-CM: 5'-AGA GTT TGA TCM TGG CTC AG-3') and a degenerate probe containing locked nucleic acids (+) (BSR65/17: 5'-56FAM-TAA +YA+C ATG +CA+A GT+C GA-BHQ1-3'). All qPCR reactions were performed in triplicate (20 µl each), with each reaction containing 0.6 µM of 27f-CM primer, 0.6 µM of 357R primer (5'-CTG CTG CCT YCC GTA G-3'), 0.25 µM of BSR65/17 probe, 10.0 µl of 2X TaqMan Environmental Master Mix 2.0 (Life Technologies, Carlsbad, CA), and 4.0 µl purified DNA. Cycling conditions were as follows: 95°C for 10 min, and 45 cycles of 94°C for 30 s, 50°C for 30 s, and 72°C for 30 s. Fluorescent readings were taken at the end of each cycle on an ABI 7500 Real-Time PCR System (Applied Biosystems, Waltham, MA). Raw amplification data were normalized to the ROX passive reference dye and analyzed with Standard Curve 3.3.0-SR2-build15 (Thermo Fisher Cloud), using automatic threshold and baseline settings. Cycle of quantification (Cq) values, defined as the average number of cycles required for normalized fluorescence to exponentially increase, were calculated. DNA derived from *Escherichia coli* ATCC 25922 containing seven 16S rRNA gene copies per genome (GenBank accession: CP009072) was quantified using a Qubit 3.0 fluorometer with a Qubit dsDNA HS Assay kit (Life Technologies, Carlsbad, CA) according to the manufacturer's instructions and used for the generation of

standard curves. To estimate qPCR efficiency, a standard curve containing seven 10-fold serial dilutions (three replicates each) ranging from 1.99 X 10^7 to 1.99 X 10^1 copies was included in each run. Prior to analyzing qPCR data with the on-line platform Thermo Fisher Cloud (Standard Curve (SR) 3.3.0-SR2-build15), an external master standard curve was generated by performing a regression of the standard curve data from all six qPCR runs. The regression values for the master external standard curve were: slope = -3.4629, y-intercept = 40.122, R^2 = 0.9798.

16S rRNA gene sequencing

16S rRNA gene sequencing was completed on an Illumina MiSeq (San Diego, CA) instrument at the University of Michigan's Center for Microbial Systems (Ann Arbor, MI). The V4 region of the 16S gene was amplified using a modified PCR approach (95° for 2 min, followed by 32 cycles of 95° for 30 s, 55° for 30 s, and 72° for 30 s, with a final elongation step at 72° for 10 min). DNA template volumes were 5 µl for urine and blank DNA extraction kits, and 3 µl for vaginal swabs. The MiSeq sequencing protocol was performed using the dual indexing sequencing strategy developed by Schloss and colleagues [219, 253].

16S rRNA gene sequence processing

Sequence data were processed using Mothur software (v1.39.5) [41]. Specifically, paired reads were assembled, quality-filtered (no ambiguous base calls, homopolymers ≤ 8 bases long), and aligned to the SILVA 16S rDNA reference database (release 102) [221, 222]. Sequences in the final dataset had an average length of 253 bp. We performed a preclustering step (diffs = 2) to reduce potential influence of sequencing errors and removed chimeras identified by UCHIME [254]. For taxonomic classification, the SILVA reference database [222] was used with a confidence threshold of 80% [255]. Sequences from an unknown domain, Eukaryota, Chloroplasts, Mitochondria, or Archaea were removed. Operational taxonomic units (OTUs)

were defined using a 3% sequence dissimilarity cutoff. Good's coverage values for all urine and vaginal samples exceeded 99%.

Statistical analyses

Bacterial culture: The rate of cultivation of bacterial phylotypes (as identified via MALDI-TOF, e.g. *Lactobacilllus*, coagulase negative *Staphylococcus*, *Ureaplasma urealyticum*) from urine was compared between Foley catheter and clean catch collection methods using generalized estimating equation models assuming a binomial distribution (i.e. detected or non-detected). Only bacterial phylotypes detected in at least 20% of the samples, regardless the method of collection, were tested. Significance of the odds ratios was assessed via Wald tests, implemented in the *geepack* package in R (v 3.4) [256]. The paired differences in the total numbers of bacterial phylotypes detected within the Foley catheter and clean catch urine samples among the women were assessed using a Poisson generalized estimating equation model.

16S rRNA gene qPCR: To assess differences in 16S rDNA abundance between each urine volume and collection method and blank DNA extraction kit controls, differences in cycle of quantification (Cq) were evaluated via Mann-Whitney tests. To assess variation in 16S rDNA abundance among urine samples of different volumes from the same women, variation in Cq values was evaluated via repeated-measures analysis of variance (ANOVA) followed by Tukey's tests for pair-wise comparisons or Friedman's ANOVA followed by Wilcoxon matched pairs tests. In component 2, differences in 16S rDNA abundance between sample types were assessed using Friedman's ANOVA followed by Wilcoxon matched pairs tests. Statistical analyses were performed using PAST software (v3.16) [202].

16S rRNA gene profile alpha and beta diversity: For Study Component 1, blank DNA extraction kit controls were sequenced twice and subsequently pooled bioinformatically. The controls with

Good's coverage values exceeding 98% were retained for analysis [n = 5, additional controls were processed during DNA extractions for Study Component 2 and used as part of the LEfSe analysis (n = 7)]. Alpha diversity in Study Component 1 was analyzed after subsampling individual libraries to 447 sequences, which corresponds to the sequence number of the least represented background technical control sample. After subsampling, Good's coverage remained above 95% for all but one sample (91%).

Alpha diversity in Study Component 2 was analyzed after subsampling individual libraries to 2007 sequences, which corresponds to the number of sequences in the second least represented sample. A sample with 445 sequences was excluded from these analyses. In Component 2, after subsampling, Good's coverage values for urine and vaginal samples remained greater than or equal to 98%.

Alpha diversity was assessed using the Chao1 index as an indicator of richness and the Shannon and Inverse Simpson indices as indicators of heterogeneity (evenness). Differences in alpha diversity between urine and background technical control samples were evaluated through t-tests or Mann-Whitney tests. For comparisons among different urine volumes (Component 1: 1 ml, 1.8 ml, 5.4 ml, 10 ml, 25 ml) or between sample types (Component 2: Foley catheter urine, clean catch urine, vaginal swabs) variation in alpha diversity was evaluated through repeated measures ANOVA followed by Tukey's matched-pairs or their non-parametric equivalents.

To evaluate differences in beta diversity of 16S rRNA gene profiles, Jaccard (i.e. composition) and Bray-Curtis (i.e. structure) similarity index values were calculated using OTU percent relative abundance data within samples and were visualized through Principal Coordinates Analyses (PCoA). Non-parametric MANOVA (NPMANOVA) tests were performed on Jaccard and Bray-Curtis similarity indices to assess differences between

background technical controls and different urine volumes (Component 1), and variation among catheter urine, clean catch urine, and vaginal samples (Component 2).

Alpha diversity indices were generated in mothur (v1.39.5) and statistically evaluated in PAST (v3.16). Beta diversity indices and PCoA plots were generated using PAST software (v3.16). Non-parametric MANOVA [257-259] tests were performed in R (version 3.4.2) with adonis in the vegan package. The "strata" parameter in adonis was used to control for repeated-measures. Heatmaps were generated via the Morpheus online tool [204].

SourceTracker analysis: SourceTracker software [121] was used to identify what percentage of OTUs found in urine samples could be attributed to contamination from vaginal samples. For each urine collection method, SourceTracker analysis was done in triplicate with a rarefaction depth of 500 and the proportions from the three model runs were averaged to give the mean percentage of OTUs predicted to be from vaginal samples. Singletons and doubletons were removed from 16S rRNA gene datasets prior to these analyses. Wilcoxon paired tests of the averaged SourceTracker runs were evaluated in PAST (v3.16).

LEfSe analysis: Linear discriminant analysis effect size, or LEfSe, was used to identify any OTUs that differed in relative abundance between each of the three biological sample types (Foley catheter urine, clean catch urine, vaginal swabs) and blank DNA extraction kits (n = 7). Singleton OTUs were removed from the datasets prior to analyses and the default settings for LEfSe were used except that the LDA score cutoff was set to 3.0.

IV. Results

Study Component 1: Comparing the bacterial load and 16S rRNA gene profiles of different volumes of urine from the same women

Patient characteristics

Eight women were included in Component 1 of the study. The median and interquartile range for age, body mass index (BMI), gestational age at sampling, and neonatal birthweight were 27.5 (25.2-28.2) years, 31.6 (28.5-45.8) kg/m^2, 39.7 (38.8-40.8) weeks, and 3,392 (3,256-3,882) grams, respectively **(Table 4.1)**.

Quantitative real-time PCR (qPCR) of 16S rRNA gene abundance in urine samples

Quantitative real-time PCR demonstrated that the median cycle of quantification (Cq) values for each volume of urine were significantly lower than the median Cq values for blank DNA extraction kit controls, regardless whether the urine was collected via Foley catheter (Mann-Whitney; $U = 0$, $p = 0.0058$) or the mid-stream clean catch method ($U = 0$, $p = 0.0081$). The volume of urine processed had an effect on Cq value for Foley catheter (repeated measures ANOVA; $F = 9.805$, $p < 0.0001$) and clean catch ($F = 28.01$, $p < 0.0001$) urine samples. For both urine collection methods, a urine volume of 5.4 ml was the lowest volume to yield Cq values that did not significantly differ from 25 ml of urine (Tukey-adjusted comparisons; **Figure 4.1**), which was the highest volume of urine investigated in this study.

16S rRNA gene profiles of urine samples

Alpha Diversity: Bacterial profile richness (Chao1 index) and heterogeneity (Shannon and Inverse Simpson indices) did not differ between any volume of catheter or clean catch urine and blank DNA extraction kits after correcting for multiple comparisons **(Table 4.2)**. A global effect of sample volume on heterogeneity was observed for Foley catheter urine; however no pairwise comparisons were significant after correcting for multiple comparisons **(Table 4.2)**.

Table 4.1. Descriptive and clinical characteristics of subjects for Study Components 1 and 2.

Study Component 1

N = 8	**Median**	**IQR[a]**
Age (yrs)	27.5	25.2 – 28.2
BMI[b] (kg/m^2)	31.6	28.5 – 45.8
GA[c] at sampling (wks)	39.7	38.8 – 40.8
Birthweight (g)	3392	3256 – 3882
Race		
African American	8 (100 %)	

Study Component 2

N = 25	**Median**	**IQR[a]**
Age (yrs)	24.0	21.0 – 29.0
BMI[b] (kg/m^2)	31.7	26.3 – 35.8
GA[c] at sampling (wks)	39.3	39.0 – 39.85
Birthweight (g)	3165	2892.5 – 3615
Race		
African American	22 (88.0 %)	
White	2 (8.0 %)	
Other	1 (4.0 %)	

[a] Interquartile range
[b] Body Mass Index
[c] Gestational Age

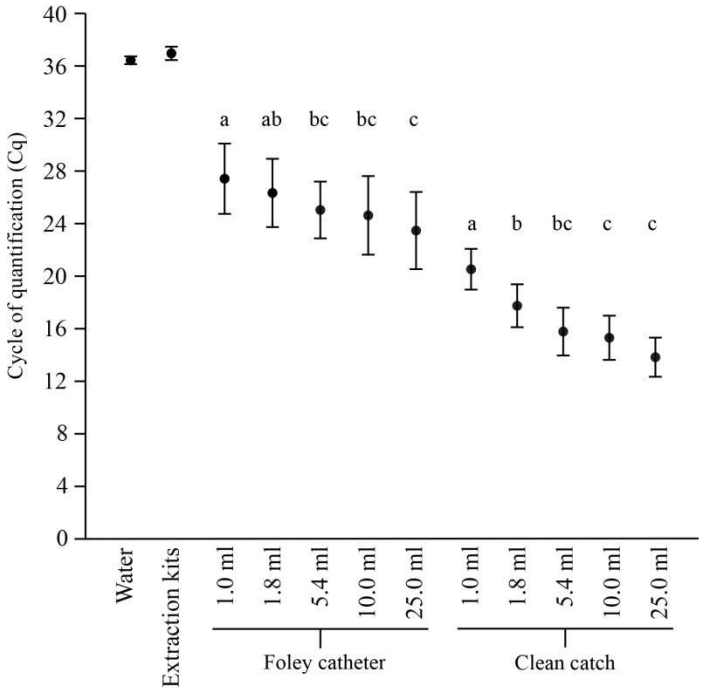

Figure 4.1. Quantitative real-time PCR (qPCR) of the 16S rRNA gene results from urine sample volumes of 1.0, 1.8, 5.4, 10, and 25 ml. Seven and six women contributed Foley catheter and clean catch urine samples, respectively. Both Foley catheter and clean catch urine collection methods yielded samples with microbial burdens exceeding those of blank DNA extraction kits (Foley catheter, N = 4, clean catch N = 5). Letters correspond to pairwise comparisons where $p > 0.05$, suggesting the microbial load in those volumes were not different from each other. Plotted values are mean ± standard error.

Table 4.2. Comparisons of alpha diversity of Foley catheter urine and clean catch urine processed at 5 different volumes and blank controls.

t-test or Mann-Whitney U test	Chao1		Shannon		Inverse Simpson	
Foley catheter v Blank controls	**t**	**P**	**t**	**P**	**t**	**P**
1.0 ml	0.2860	0.7814	1.6081	0.1423	1.5260	0.1613
1.8 ml	1.4256	0.1877	1.3154	0.2209	1.1052	0.2978
5.4 ml	2.0131	0.0746	14*	0.9247	0.8111	0.4382
10 ml	2.4313	0.0379	1.9453	0.0836	1.6474	0.1339
25 ml	1.4779	0.1736	1.4726	0.1750	1.5136	0.1644
Clean catch v Blank controls	**t**	**P**	**t**	**P**	**t**	**P**
1.0 ml	0.4407	0.6688	2.1093	0.0611	1.7277	0.1147
1.8 ml	0.0219	0.9830	2.2362	0.0493	1.7596	0.1090
5.4 ml	0.7251	0.4850	2.3172	0.0430	1.8545	0.0934
10 ml	0.2532	0.8053	2.5737	0.0277	2.0930	0.0628
25 ml	0.0377	0.9707	2.8564	0.0171	2.1924	0.0531
Repeated-measures ANOVA or Friedman's ANOVA	**F**	**P**	**F**	**P**	**F**	**P**
Foley catheter						
Volume	1.7700	0.1998	10.45**	0.0349	4.215	0.0233
Pairwise comparisons				All > 0.01		All > 0.01
Clean catch	**F**	**P**	**F**	**P**	**F**	**P**
Volume	0.5478	0.7033	1.462	0.2600	1.911	0.1578
Pairwise comparisons						

***Mann-Whitney U; **chi^2**

Beta Diversity: The composition (Jaccard index) and structure (Bray Curtis index) of the bacterial profiles of all five volumes of Foley catheter and clean catch urine samples differed from those of blank DNA extraction kit controls **(Table 4.3, Table 4.4, Figure 4.2, Figure 4.3)**. Subject identity, not urine sample volume, principally influenced the composition and structure of urine bacterial profiles, regardless the method of collection **(Table 4.3, Table 4.4**, **Figure 4.3)**.

Study Component 1 Outcome

Given that a sample volume of 5.4 ml was the lowest volume of urine to yield Cq values that did not differ from those of 25 ml of urine, regardless collection method, a urine sample volume of 5.4 ml was used in Component 2 of the study.

Study Component 2: Evaluating differences in microbial burden and 16S rRNA gene profiles between Foley catheter and mid-stream clean catch urine in relation to those of vaginal swabs

Patient characteristics

Table 4.1 describes the demographic and clinical characteristics of the 25 women included in Component 2 of the study [the median and interquartile range for age were 24 (21-29) years, for body mass index were 31.7 (26.3-35.8) kg/m^2, for gestational age were 39.3 (39-39.85) weeks, and for neonatal birthweight were 3,165 (2,892.5-3,615) grams]. Twenty-two women were African-American, two were Caucasian, and one was self-reported as Other. Seven women (28%) had a history of at least one lifetime UTI, and two (8%) experienced a UTI episode earlier during this pregnancy. However, none had a UTI within 30 days of sampling/delivery.

Table 4.3. Statistical analysis of bacterial community composition (Jaccard similarity index) for Foley catheter urine and clean catch urine processed at five different volumes and compared to blank controls.

Jaccard NPMANOVA	F	R^2	P
Foley catheter v Blank controls			
1.0 ml	1.0893	0.0901	0.0056
1.8 ml	1.1143	0.0920	0.0046
5.4 ml	1.1300	0.0932	0.0019
10 ml	1.2040	0.0912	0.0008
25 ml	1.1190	0.0923	0.0086
Clean catch v Blank controls			
1.0 ml	1.1223	0.0926	0.0026
1.8 ml	1.1113	0.1000	0.0031
5.4 ml	1.1021	0.0993	0.0041
10 ml	1.0916	0.0903	0.0062
25 ml	1.0863	0.0980	0.0095
Foley catheter			
Subject (n = 4)	1.5412	0.2241	0.0001
Volume	1.0023	0.1943	0.4393
Clean catch			
Subject (n = 5)	1.7169	0.2555	0.0001
Volume	1.0028	0.1492	0.4313

Table 4.4. Statistical analysis of bacterial community structure (Bray-Curtis similarity index) for Foley catheter urine and clean catch urine processed at five different volumes and compared to blank controls.

Bray-Curtis NPMANOVA	F	R^2	P
Foley catheter v Blank controls			
1.0 ml	2.7119	0.1978	0.0011
1.8 ml	2.5168	0.1862	0.0011
5.4 ml	2.4866	0.1844	0.0009
10 ml	3.4612	0.2239	0.0008
25 ml	2.4820	0.1841	0.0011
Clean catch v Blank controls			
1.0 ml	2.2496	0.1698	0.0012
1.8 ml	1.6062	0.1384	0.0321
5.4 ml	1.6077	0.1385	0.0299
10 ml	2.3175	0.1740	0.0020
25 ml	1.6200	0.1394	0.0353
Foley catheter			
Subject (n = 4)	12.5868	0.7004	0.0001
Volume	1.0389	0.0771	0.4032

Clean catch			
Subject (n = 5)	27.4836	0.8427	0.0001
Volume	1.1295	0.0346	0.3222

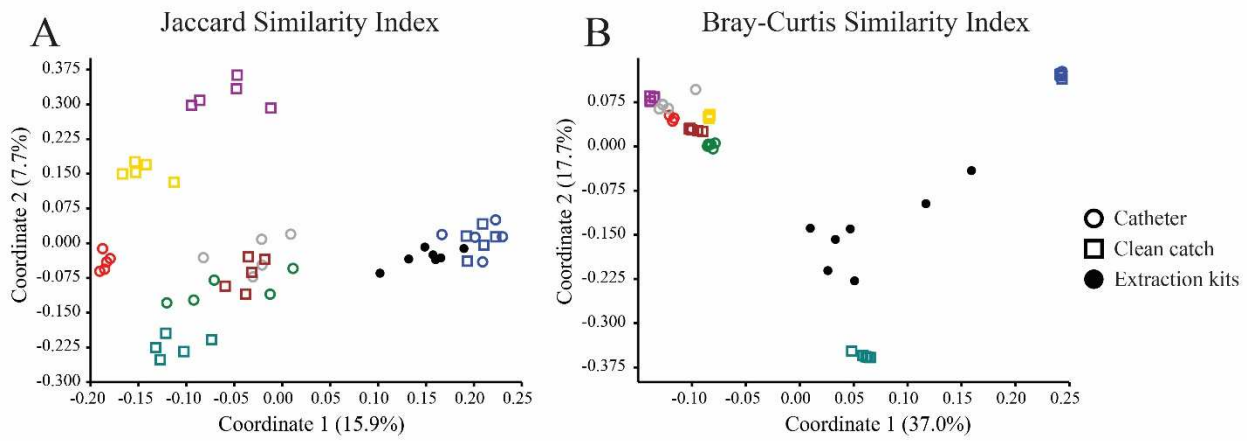

Figure 4.2. Principal Coordinates Analysis (PCoA) plots: using A) Jaccard and B) Bray-Curtis similarity indices illustrating that the composition and structure of the bacterial profiles of all urine samples, independent of sample volume or collection method, were distinct from those of DNA extraction kit controls. Subject identity, indicated by color, was the principal driver of urine bacterial profiles.

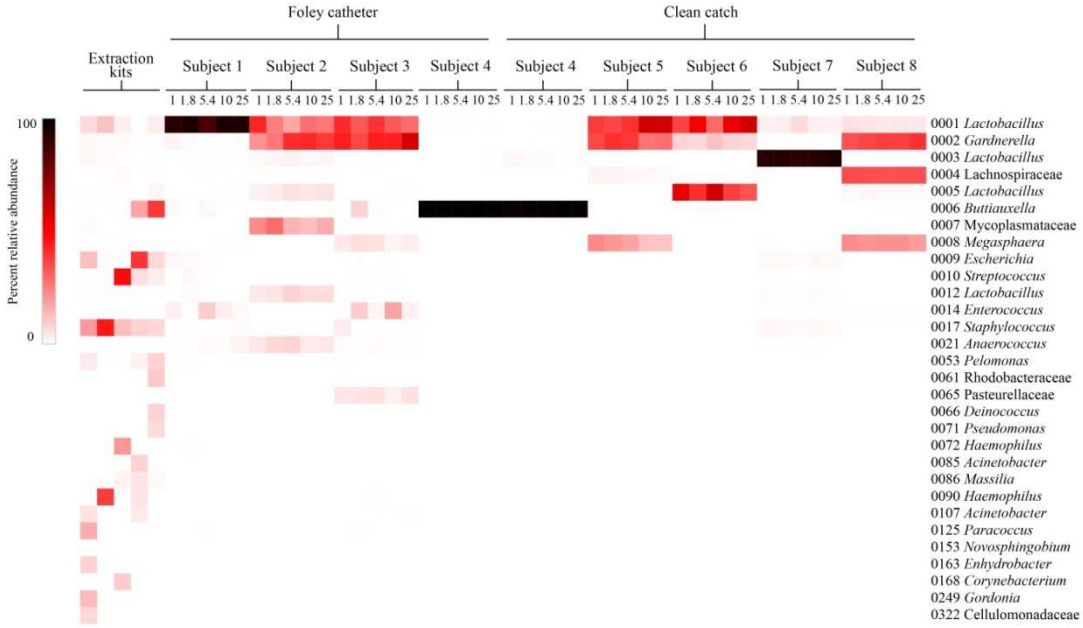

Figure 4.3. Heatmap illustrating variation in the profiles of prominent OTUs (≥1% average relative abundance) among urine samples from subjects, ordered by urine collection method and sample volume. Urine volume had little influence on bacterial profiles, while subject identity was the primary driver. The bacterial profiles of blank DNA extraction kits are distinct from urine and are indicated on the left.

Bacterial cultivation

Only three types of bacteria (*Lactobacilllus* species, coagulase negative *Staphylococcus* species, and *Ureaplasma urealyticum*) were cultured from at least 20% of all urine samples. Each was cultured less frequently from urine obtained through a Foley catheter than through mid-stream clean catch urine **(Table 4.5; Figure 4.4)**. On average, urine collected with a Foley catheter yielded three less types of bacteria than paired urine samples collected through clean catch (Poisson generalized estimating equations model; $p < 0.001$). *Ureaplasma urealyticum* was the most frequently detected bacteria in Foley catheter urine (13/25 women).

Quantitative real-time PCR (qPCR) of 16S rRNA gene abundance in urine and vaginal swab samples

The bacterial load of clean catch urine exceeded that of catheter urine (**Figure 4.5**; Wilcoxon matched-pairs: $W = 325$, $p < 0.0001$). The relationship was the same for vaginal swabs and catheter urine (**Figure 4.5**; $W = 321$, $p < 0.0001$).

16S rRNA gene profiles of paired catheter urine, clean catch urine, and vaginal swab samples

Alpha Diversity: Catheter urine, clean catch urine, and vaginal swab samples did not vary in richness (Chao1 index; Friedman's ANOVA: $p > 0.05$), but they did vary in heterogeneity based on Shannon ($Chi^2 = 7.28$, $p = 0.027$) and Inverse Simpson ($Chi^2 = 7.44$, $p = 0.025$) indices **(Figure 4.6A & B)**. The bacterial profiles of Foley catheter and clean catch urine were more heterogeneous than those of vaginal swabs (Wilcoxon matched pairs with Bonferroni corrections applied: Foley catheter, Shannon index: $W = 266$, $p = 0.0054$, Inverse Simpson index: $W = 257$, $p = 0.011$; clean catch, Shannon index: $W = 277$, $p = 0.0021$, Inverse Simpson index: $W = 253$, $p = 0.015$). The heterogeneity of catheter and clean catch urine bacterial profiles did not differ ($p > 0.05$).

Table 4.5. Odds ratios of detecting bacterial phylotypes through culture in urine obtained from a Foley catheter compared to urine collected through the mid-stream clean catch method.

Bacterial phylotype identified by MALDI-TOF	Odds Ratio	p	q
***Lactobacilllus* species**	0.11	0.000	0.000
Coagulase negative *Staphylococcus* species	0.03	0.000	0.000
Ureaplasma urealyticum	0.34	0.007	0.007

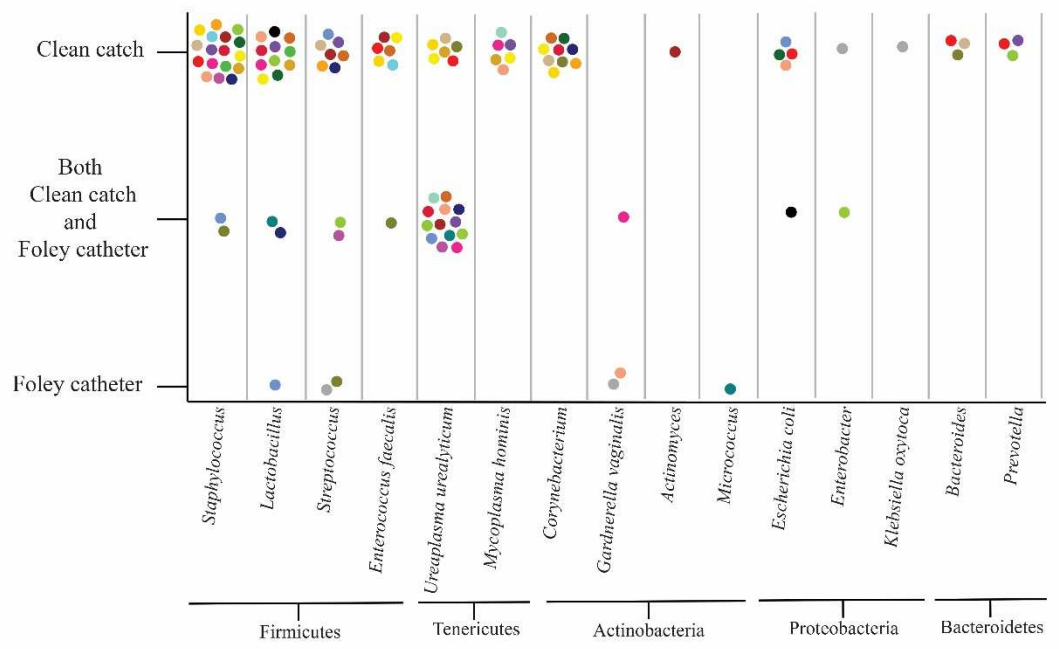

Figure 4.4. Urine bacterial cultivation results indicating differential recovery of bacterial phylotypes from catheter urine, clean catch urine, or both. Subject identity is indicated by color. In all but two occurrences, *Staphylococcus* species recovered were coagulase negative, the exception being *S. aureus* recovered in both urine samples of one patient and the clean catch urine of the other patient.

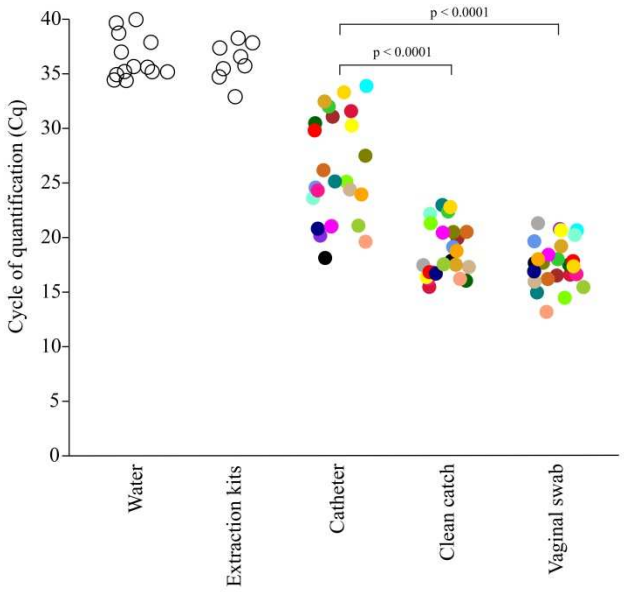

Figure 4.5. Quantitative real-time PCR (qPCR) of the 16S rRNA gene results from Foley catheter urine, clean catch urine, and vaginal swabs. 25 women contributed paired Foley catheter urine, clean catch urine, and vaginal swab samples. Color represents subject identity, Cycle of quantification (Cq) values were averaged over multiple runs.

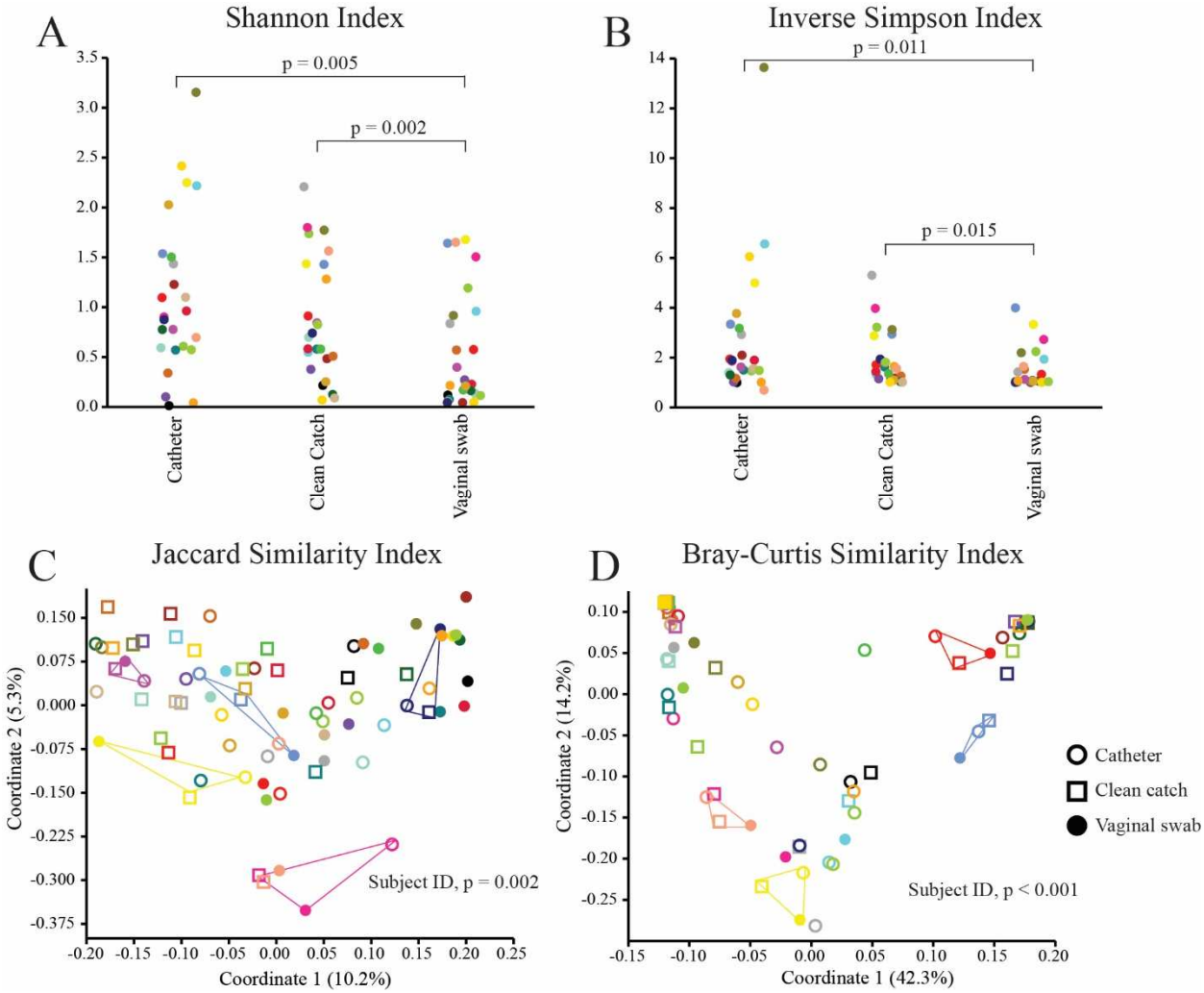

Figure 4.6. Jitter and Principal Coordinates Analysis (PCoA) plots illustrating alpha and beta diversities of Foley catheter urine, clean catch urine, and vaginal swabs collected from the same women. Panels A and B show differences in heterogeneity between sample types, with catheter urine having the greatest diversity for both indices. Panels C and D illustrate the composition and structure of the bacterial profiles of the three sample types. Several subjects are highlighted to illustrate the influence of individual identity on the bacterial profiles. Subject identity is indicated by the same color scheme across the 4 panels.

Beta Diversity: Overall, subject identity was the principal driver of the composition and structure of the bacterial profiles of Foley catheter urine, clean catch urine, and vaginal swabs **(Table 4.6, Table 4.7, Figure 4.6C & D, Figure 4.7)**. Nevertheless, controlled for subject identity, the composition of the three sample types differed from one another **(Table 4.6, Figure 4.6C)**. With respect to structure, the bacterial profiles of catheter urine differed from those of clean catch urine and vaginal swabs, but the profiles of clean catch urine did not differ from the profiles of their paired vaginal swabs **(Table 4.7, Figure 4.6D)**.

SourceTracker analyses: SourceTracker analysis found there was a greater contribution of OTUs explained by vaginal swabs in clean catch urine than in catheter urine (**Figure 4.8**; Wilcoxon paired test: $W = 279$, $p < 0.001$).

Bacterial profiles in detail: The bacterial profiles of catheter urine, clean catch urine, and vaginal swabs were dominated by *Lactobacillus* and *Gardnerella* **(Figure 4.7)**. BLAST analyses indicated that OTUs 1 and 2 were *Lactobacillus iners* and *Gardnerella vaginalis*, respectively. OTUs 3, 5, and 12 were each identified as multiple species of *Lactobacillus.* OTU 3 was identified as *Lactobacillus crispatus* (most likely [48, 68, 115, 118, 133, 169])*, acidophilus,* or *gallinarum,* OTU 5 was identified as *Lactobacillus jensenii* (most likely [48, 68, 115, 118, 133, 169]) or *fornicalis*, and OTU 12 was identified as *Lactobacillus gasseri* (most likely [48, 68, 115, 118, 133, 169]) or *johnsonii.* Thus, consistent with prior observations [50, 51, 68, 133, 227], the urogenital bacterial profiles of pregnant women were largely comprised of three community state types: 1) dominance by *Lactobacillus crispatus*; 2) dominance by *Lactobacillus iners*; or, 3) co-dominance by *Lactobacillus iners* and *Gardnerella vaginalis.* Both *Lactobacillus* and *Gardnerella* were rarely detected in the cultivation surveys of urine, suggesting their molecular signals may have been due to contamination, although both genera often require specialized

Table 4.6. Statistical analysis of bacterial community composition (Jaccard similarity index) for Foley catheter urine, clean catch urine, and vaginal swabs.

Jaccard NPMANOVA	F	R^2	P
Foley Catheter v Clean Catch v Vaginal Swab			
Subject (n = 25)	1.02349	0.32957	0.0019
Sample Type	0.98388	0.02640	0.7941
Foley Catheter v Clean Catch			
Sample Type within subject	1.0515	0.02144	0.0022
Clean Catch v Vaginal Swab			
Sample Type within subject	1.1345	0.02309	0.0014
Foley Catheter v Vaginal Swab			
Sample Type within subject	1.1306	0.02301	0.0027

Table 4.7. Statistical analysis of community structure (Bray-Curtis similarity index) for Foley catheter urine, clean catch urine, and vaginal swabs.

Bray-Curtis NPMANOVA	F	R^2	P
Foley Catheter v Clean Catch v Vaginal Swab			
Subject (n = 25)	1.55064	0.42797	0.0002
Sample Type	0.87084	0.02003	0.6228
Foley Catheter v Clean Catch			
Sample Type within subject	0.88327	0.01807	0.0222
Clean Catch v Vaginal Swab			
Sample Type within subject	0.92679	0.01894	0.4834
Foley Catheter v Vaginal Swab			
Sample Type within subject	1.9757	0.03953	0.0460

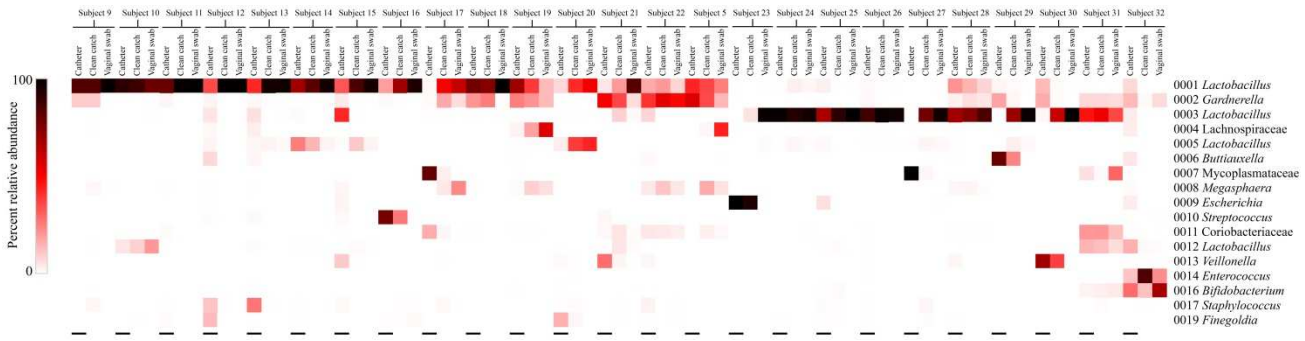

Figure 4.7. Heatmap illustrating variation in the profiles of prominent OTUs (≥1% average relative abundance) among paired Foley catheter urine, clean catch urine, and vaginal swab samples from 25 pregnant subjects. The order was determined by hierarchical clustering of Bray-Curtis similarity indices of clean catch samples. Bars along the bottom horizontal axis mark the catheter sample of each patient.

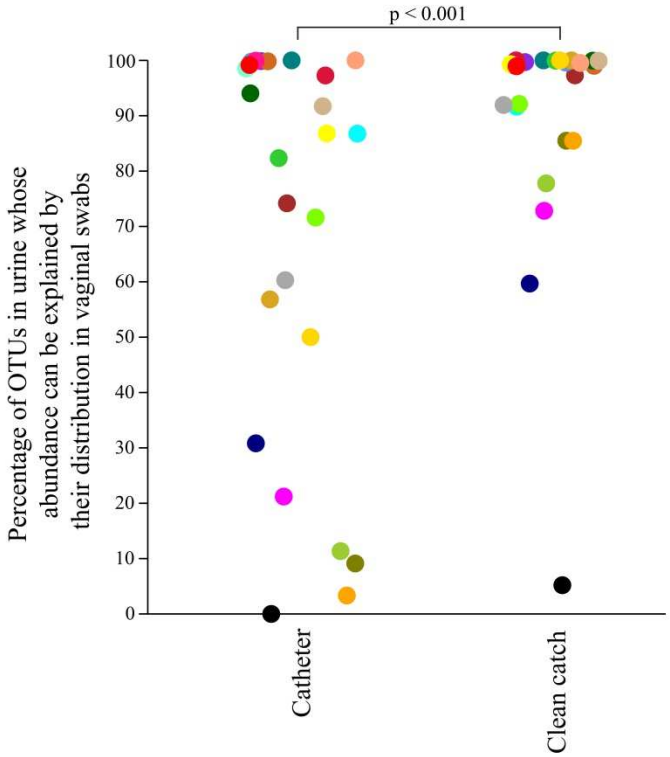

Figure 4.8. SourceTracker analysis comparing the percentage of OTUs explained by vaginal swabs among Foley catheter urine and clean catch urine. Colors represent subject identity.

media even when recovering them from pure culture. Besides *Lactobacillus* and *Gardnerella,* the bacterial profiles of catheter urine also variably contained a high relative abundance (at least one sample with >50% relative abundance) of 5 additional OTUs: an unclassified Mycoplasmataceae (OTU7), *Escherichia* (OTU 9), *Buttiauxella* (OTU 9)*, Streptococcus* (OTU 10)*,* and *Veillonella* (OTU 13). Notably, 5 additional OTUs, not previously discussed, were detected in at least 12 catheter urine samples and had a relative abundance of ≥ 10% in at least one sample. These OTUs included an unclassified Coriobacteriaceae (OTU 11), *Staphylococcus* (OTU 17), *Finegoldia* (OTU 19), *Ureaplasma* (OTU 26), and *Peptoniphilus* (OTU 27).

LEfSe analysis: Linear discriminant analysis effect size (LEfSe) analyses identified eight OTUs that were consistently more abundant in controls than any of the biological sample types, suggesting these OTUs are likely contaminants **(Figure 4.9)**. These OTUs were identified as *Escherichia* (OTU 9), *Staphylococcus* (OTU 17), *Pelomonas* (OTU 53), *Massilia* (OTU 86), *Haemophilus* (OTU 90), *Virgibacillus* (OTU 102), *Acinetobacter* (OTU 107), *Cloacibacterium* (OTU 329). Analyses comparing catheter urine to vaginal swabs identified 17 OTUs more relatively abundant in catheter urine **(Figure 4.10)**. While this analysis did not control for patient identity, four of these seventeen OTUs were also more relatively abundant in catheter urine than negative controls (**Figure 4.10**, highlighted in blue). These included *Finegoldia* (OTU 19), *Ureaplasma* (OTU 26), *Anaerococcus* (OTU 49), and an unclassified Clostridiales [OTU 43 (BLAST query identified it as *Fenollaria massiliensis*)].

After identifying these four OTUs as potentially indicative as members of a bladder microbiota (i.e. more relatively abundant in catheter urine than in controls, and possibly more than vaginal samples as well), their abundances in catheter urine and vaginal swabs were compared directly through paired testing. All four OTUs were more abundant in catheter urine

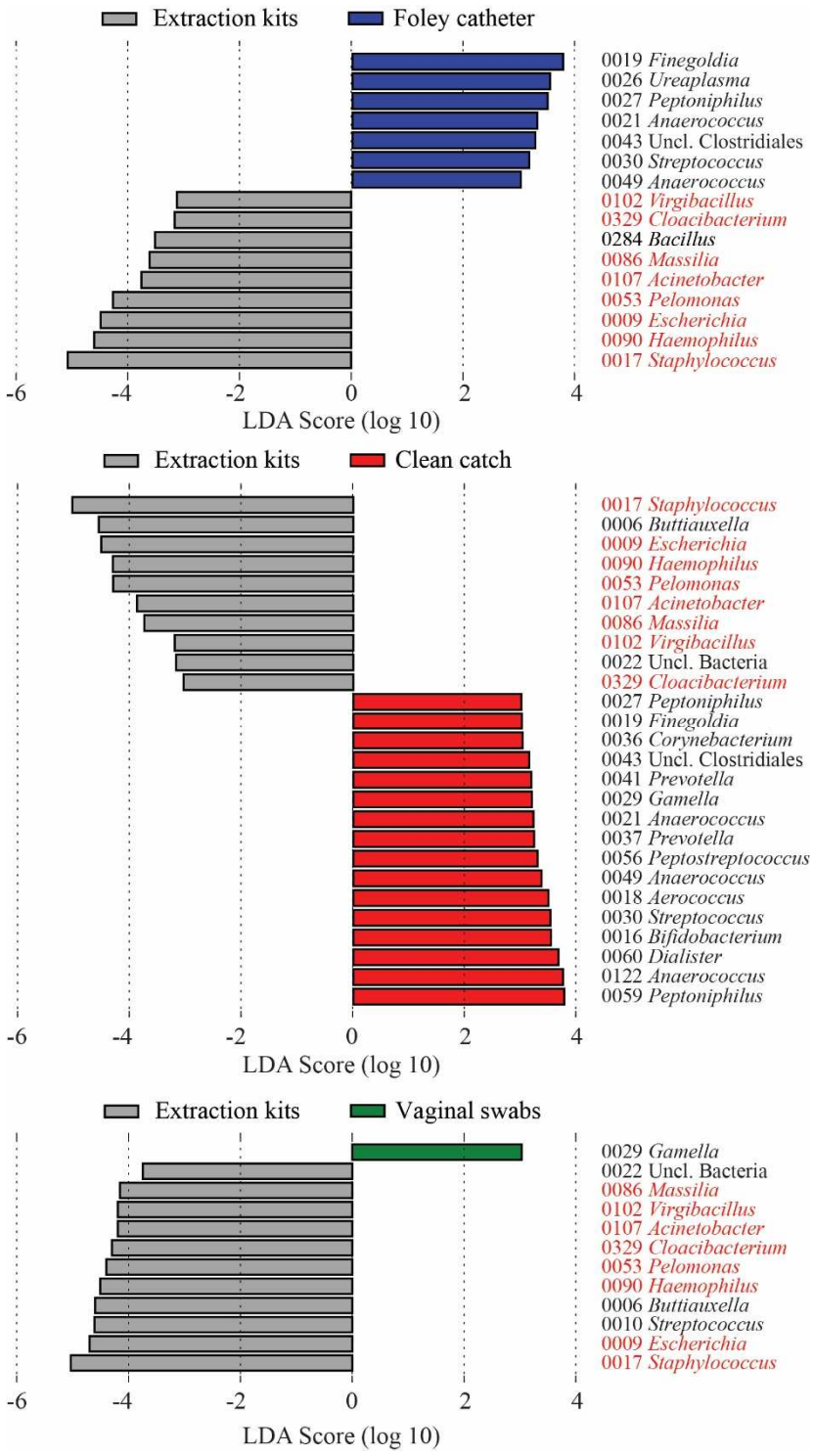

Figure 4.9. Linear discriminant analysis Effect Size analyses identified several bacteria that were more relatively abundant in blank extraction kits. Analyses of 16S rRNA gene sequence datasets from DNA extraction compared to each sample type. OTUs highlighted in red were more relatively abundant in extractions kits than in all 3 biological samples suggesting them as likely contaminant sequences.

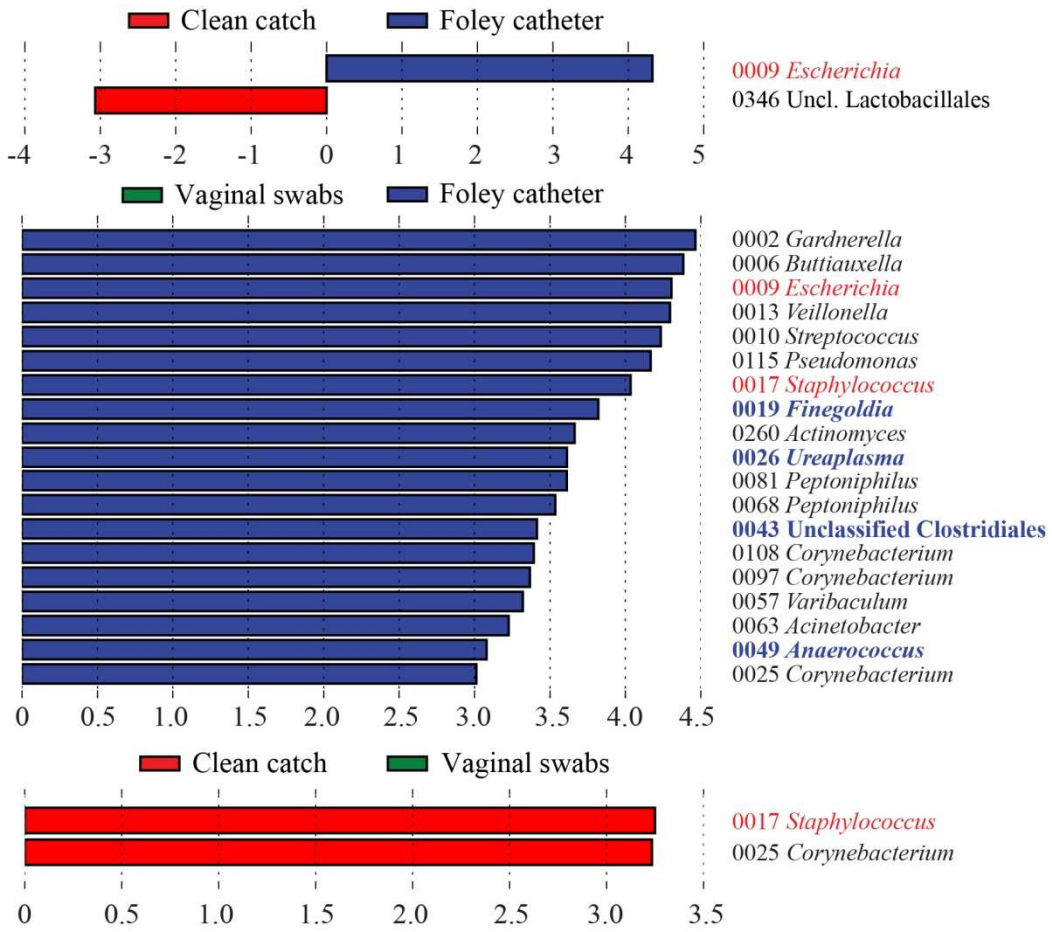

Figure 4.10. Linear discriminant analysis Effect Size analyses identified several bacteria that were more relatively abundant in Foley catheter urine over vaginal samples, suggesting they are members of a urine and bladder microbiota. Analyses of sequence datasets from Foley catheter urine with DNA extraction kits, clean catch urine, and vaginal swabs. Highlighted in blue are taxa identified as being more relatively abundant in catheter urine than in both extraction kit controls and vaginal swabs.

than vaginal swabs with sequential Bonferroni corrections applied (Wilcoxon matched pairs: OTU 19, W = 152, p = 0.02; OTU 26, W = 134, p = 0.006; OTU 43, W = 69, p = 0.019; OTU 49, W = 74, p = 0.046). Notably, vaginal swab samples did not have any OTUs that were more relatively abundant than they were in Foley catheter urine samples, which can be explained because the vaginal samples were less diverse overall – they were dominated by a few OTUs, accounting for over 90% of the average OTU abundance, and these OTUs were also identified in the urine samples (OTU 1 *Lactobacillus*, OTU 2 *Gardnerella*, OTU 3 *Lactobacillus*, OTU 4 unclassified *Lachnospiraceae*, OTU 5 *Lactobacillus*, and OTU 16 *Bifidobacterium*).

V. Discussion

Principal findings of the study

(1) Quantitative real-time PCR showed that the bacterial load of urine exceeded technical controls regardless of the urine collection method (Foley catheter or clean catch) or the volume of urine processed; (2) A urine volume of 5.4 ml was the lowest to yield a similar 16S rRNA gene load and profile as 25 ml of urine, which was the largest urine volume we investigated; (3) Via cultivation, three isolates were detected in at least 20% of all samples (*Lactobacilllus* species, *Staphylococcus* species coagulase negative, and *Ureaplasma urealyticum*) and all three were detected less frequently in catheter than clean catch urine; (4) An average of three less isolates were recovered from catheter urine clean catch urine; (5) Molecular sequencing techniques showed that the bacterial profiles of clean catch urine were more similar to vaginal swabs than catheter urine, suggesting a greater influence of vulvovaginal contamination on clean catch samples; (6) *Lactobacillus iners, L. crispatus,* and *Gardnerella vaginalis* were the most relatively abundant bacteria among all 3 sample types; (7) *Ureaplasma* (detected in culture and molecular surveys) and multiple Gram-positive species including elusive Gram-positive

anaerobic cocci (GPAC) such as *Finegoldia* and *Anaerococcus* were more relatively abundant in in catheter urine than clean catch urine or vaginal swabs.

Overall, our study suggests that resident bacterial communities exist in the bladder and urine of pregnant women, and that while there is overlap with the vaginal microbiota, there are also distinct lower abundance taxa in the bladder.

Urinary Tract Infections (UTIs) and Pregnancy

Approximately 10% of women report having at least one episode of UTI in the previous 12 months and the lifetime probability of a woman having at least one UTI event is around 60% [230, 260-262]. Among women with culture-confirmed UTIs, around 28% had recurrence within six months [230, 263]. The bacterium most responsible for UTIs is *Escherichia coli*, followed by *Staphylococcus saprophyticus*, *Streptococcus agalactiae*, and *Klebsiella* and *Enterococcus* species [158, 233, 244, 249]. The occurrence of a UTI during pregnancy is associated with significant odds ratios of 1.4 for low birth weight, 1.3 for preterm birth (< 37 weeks), 1.4 for maternal hypertension and preeclampsia, 1.6 for maternal anemia (hematocrit less than 30%), and 1.4 for chorioamnionitis [157, 158]. The risk of occurrence of these complications is higher among patients with pyelonephritis [158, 264]. Acute pyelonephritis occurs in 2% of pregnant women and recurs in 33% of them [134, 158].

In 2017, a European interdisciplinary group consisting of 17 representatives of 12 medical societies was formed to update the diagnosis and management of uncomplicated UTIs in non-pregnant premenopausal women and concluded that the diagnosis of uncomplicated cystitis be based on clinical criteria evaluating the symptoms of the patient and the course of the disease [265]. They also recommended that no microbiological examination is needed in asymptomatic non-pregnant patients [265]. In pregnancy, asymptomatic bacteriuria is to be detected and treated

because of the potential serious complications, including pyelonephritis in 30-40 % of the cases [249]. The American college of Obstetricians and Gynecologists (ACOG) and other societies recommend urine culture as one of the routine tests to be obtained early in pregnancy [266, 267], with mid-stream clean catch being the most common collection approach. If mid-stream clean catch urine culture is performed, the detection of *Escherichia coli* is predictive of bacterial UTI while the detection of other bacteria such as enterococci and group B streptococci is not predictive [265]. Negative urine dipsticks and urine microscopy are useful to rule out the likelihood of presence of UTI in asymptomatic non-pregnant women [265]. Urine culture is useful if only one type of bacteria was cultivated [265].

Vulvovaginal contamination of urine samples

Prior culture-based studies concluded that clean catch urine samples obtained from female patients have poor ability to detect UTI because of contamination from the skin and vaginal microbiota [268-270]. A study of 113 asymptomatic pregnant women showed a high level of contamination in clean catch samples [268]. Baerheim et al. [269] found that employing precautions such as cleaning the perineum or obtaining mid-stream samples led to similar contamination rates as obtaining samples without any precautions [269]. Lifshitz et al. [270] evaluated 242 symptomatic female patients divided into 3 groups 1) no cleaning, 2) perineal cleaning and midstream sampling, and 3) perineal cleaning, midstream sampling and vaginal tampon. Contamination rates in the three groups were all similar at approximately 30% [270].

In women undergoing cesarean deliveries, two randomized clinical trials showed an increase in the incidence of UTIs in the indwelling bladder catheterization group compared to the no catheterization group [271-273]. Mid-stream clean catch remains a more convenient approach for patients and health care staff; however, culture results of clean catch samples need careful

interpretation since around 30 percent of clean catch samples showed contamination in previous studies [268-270]. Our cultivation studies confirmed that Foley catheter samples yield less cultivars than clean catch samples and are less likely to be contaminated by vaginal microbiota than are clean catch samples.

In the context of next-generation sequencing, it is important that: 1) vulvovaginal contamination be limited during urine sample collection; 2) collected urine specimens be promptly frozen to mitigate growth and replication of acquired contaminants; 3) a sufficient volume of urine is collected for effective DNA extraction; and 4) that background technical controls be included to account for potential DNA contamination.

5.4 ml of urine is an appropriate volume for 16S rRNA sequencing

Previous urine microbiota studies have utilized different sample volumes of urine for 16S rRNA gene analysis [69, 114, 117, 118, 226, 227]. Component 1 of our study showed that a volume of 5.4 ml represents an appropriate volume for molecular survey analysis. While a bacterial signal was obtained from all volumes (1, 1.8, 5.4, 10 and 25 ml), there were higher bacterial loads detected (lower Cq) in the 5.4, 10.0, and 25.0 ml groups compared to the 1.0 and 1.8 ml groups. There was no statistical significance between 5.4 ml and either the 10.0 or 25.0 ml groups. These results show that 5.4 ml urine volume yields an appropriate quantity of DNA for 16S rRNA gene sequencing and that the DNA yield would not substantively change with a greater volume of urine processed (at least up to 25 ml). Clean catch urine samples consistently had a greater bacterial load than catheter urine samples indicating that the biological source of the sample (clean catch versus Foley catheter) has a greater influence on bacterial load than the volume of the sample.

Both urine collection methods resulted in bacterial profiles that were significantly different than controls regardless of the urine volume processed. While urine samples primarily exhibited dominance of 2 to 3 OTUs, controls showed a more even spread between various OTUs including common contaminants such as *Escherichia*, *Streptococcus* and *Staphylococcus*.

Our results suggest that a bladder microbiota exists during pregnancy

For Study Component 2, while individual identity was the primary influence on 16S rRNA gene profiles, differences were found for both beta diversity indices (i.e. Jaccard and Bray Curtis) between the three sample types: catheter urine, clean catch urine, and vaginal swabs. Pairwise comparisons showed that the bacterial profiles of catheter urine were different than those of both clean catch urine and vaginal swabs in composition and structure, whereas the profiles of clean catch urine were different from vaginal swabs only in structure. This suggests that urine collected via the clean catch method has a bacterial profile more similar to that of the vagina due to proximity and increased likelihood of vulvovaginal contamination. In this way, clean catch urine likely represents a dynamic mix of members of the skin and vaginal microbial communities in addition to those of a resident bladder microbial community. Whilst urine collected via a catheter is still likely subject to some degree of vulvovaginal contamination, our data suggest a microbial community in the bladder is distinct in several lesser abundant microbes.

LEfSe analyses identified seven OTUs that were more relatively abundant in Foley catheter urine than in blank extraction controls, and four of these OTUs were also more relatively abundant in catheter urine than in vaginal swab samples, suggesting that these OTUs may be representative of members of the bladder microbiota, at least in some women. While only *Ureaplasma* was recovered from cultivation surveys, *Finegoldia* and *Anaerococcus* species are

Gram-positive anaerobic cocci recalcitrant to culture, typically requiring long incubation times and complex growth requirements [274]. This may explain why they were not recovered in culture. Other groups investigating urine via molecular surveys have also identified *Finegoldia* and *Anaerococcus* species in non-pregnant females, in healthy cohorts as well as among women with non-UTI urinary disorders [48, 67, 116, 118, 153, 275]. The fourth OTU, while originally described as an unclassified Clostridiales, was a 100% sequence match to the newly identified genus and species *Fenollaria massiliensis* gen. nov., sp. nov [276, 277]. Its relatively recent discovery may explain why it was not detected by culture as most clinical microbiology laboratories are likely unfamiliar with this organism. This microorganism is discussed in more detail below.

Ureaplasma species, like other members of the class Mollicutes, lack a cell wall, require specialized media, and can require long incubation times [248, 278]. *Ureaplasma* is well-documented in its association with diseases of the urinary tract [48, 63, 64, 67, 278, 279] as well as adverse pregnancy outcomes, including preterm birth [49, 140, 280-283] and diseases of the neonate [284-287]. In fact, it is the most common microorganism found in the amniotic cavity [140, 288]. Despite its associations with disease, *Ureaplasma* is frequently detected by culture and molecular surveys in urine of asymptomatic women, pregnant [65, 137, 156, 169, 246, 248] and non-pregnant [48, 67, 68, 116], suggesting its potential role in a bladder microbiota. The various consequences associated with the presence of *Ureaplasma* in the bladder of pregnant women is likely multifactorial, of which adverse outcomes are likely associated with a combination of the individual's own immune response [289], the composition and structure of the broader microbial community [69, 171], and the specific strain of *Ureaplasma [290-293]*.

Finegoldia has been associated with the genitourinary tract, gastrointestinal tract, and skin as a commensal but has also been isolated from and attributed to infections from wounds and various body sites making it an opportunistic pathogen. Difficulty in cultivation has been evident in clinical reports where accurate diagnoses were dependent on detection via PCR despite cultures yielding negative results [294-296].

Anaerococcus species are also commensals of the skin, gastrointestinal tract, and oral cavity and members have been isolated from vaginal secretions and purulent wounds [274, 297, 298]. Literature reports successful growth of *Anaerococcus* species on standard anaerobic plate types [299, 300] by some groups, while others indicate the addition of supplemental nutrients such as hemin and vitamin K [274, 299-302]. Reported incubation times have also varied from 2 days up to 7 days [300, 302].

Fenollaria massiliensis is a newly discovered and understudied anaerobic rod recovered and characterized from osteoarticular, genital, and tissue samples, and is suggested to be a genital-associated microbe [276, 277]. The two studies describing this organism report growth on several enriched media types after 72 hours and on supplemented Brucella Blood Agar after 48 hours under anaerobic conditions.

Our study showed that catheter urine samples do yield a 16S rRNA gene signal beyond that evident in controls and suggest catheterization may be an appropriate sampling method for evaluating any microbial communities that may exist in the bladder. Also, our results suggest that the vaginal microbiota influences or contaminates clean catch urine to a larger degree than catheter urine, and that while some catheter samples are still influenced by vaginal microbes, a potentially unique signal may exist in some individuals. Our evidence suggests that underlying *Lactobacillus* and *Gardnerella* abundance in the urine of pregnant women, anaerobic organisms

like *Finegoldia, Anaerococcus*, and *Peptoniphilus* may be low abundant members of a bladder microbial community. However, to more confidently assert that the bladder contains microbiota in pregnancy, a suprapubic sampling approach would provide better insight. We suspect that the gold standard for investigating a potential bladder microbiota would be suprapubic aspiration of urine in concert with tissue sampling of the bladder epithelium. Culture should be performed to discriminate live bacteria from remnant DNA from dead bacteria.

Strengths and limitations

This is the first extensive study that attempts to characterize the urinary microbiota in pregnancy by comparing Foley catheter, clean catch, and vaginal samples from 25 women. In addition, this is the first study that compared different volumes of urine to determine the optimal volume for performing 16S rRNA gene surveys. Furthermore, this study utilized cultivation, qPCR, and sequencing approaches to study the existence and viability of microbiota in the bladder. The main limitation of the study is that our population mainly consists of one ethnic group (i.e. African American). It is possible that other ethnic groups may have a different bladder microbiota. Non-pregnant women were not included in this study, therefore differences and similarities between a healthy female bladder microbiota and a pregnant female bladder microbiota cannot be addressed. Additionally, is difficult to assess if and how much of the bacterial signal in catheter urine was due to vulvovaginal contamination, specifically regarding the top 3 most abundant taxa. More extensive culture methods may have allowed the lower abundance anaerobic organisms to be recovered.

Conclusions

Our study suggests that resident bacterial communities exist in the bladder and urine, and that there is overlap with the vaginal microbiota. While the most frequent microorganisms

recovered by Foley catheter samples were *Ureaplasma,* molecular surveys identified low abundance anaerobic bacteria in addition to *Ureaplasma* as potential members of a bladder microbiota.

Future directions

Future research should endeavor to evaluate the typical presence of a microbial community within the urinary tract of pregnant women by comparing and analyzing 16S rRNA gene sequence data from pregnant women delivering preterm (condition; defined as delivering ≤ 37 weeks) and at term (biological control; defined as delivering > 37 weeks) and the appropriate technical controls. In doing so, we can assess influences of resident microbiota on perinatal health and pregnancy outcomes and identify bacteria or bacterial communities whose presence or absence can serve as potential risk indicators. Ultimately, being able categorize and describe bladder microbial communities and their associations with preterm birth should lead to potential targets, therapeutic interventions, and other methods for treating or modifying the bladder microbiota and lessening the impact of urinary tract-associated pregnancy complications.

CHAPTER 5: GENERAL CONCLUSIONS

Contemporary sequencing technologies have allowed for the examination, in some cases re-examination, and deep characterization of microorganisms in a multitude of environments, from the seemingly inhospitable (e.g., deep sea hydrothermal vents and Antarctic permafrost) to the clearly hospitable and heavily populated (e.g., sea water and the human gut). The use of 16S rRNA gene and metagenomic sequencing has become so widespread that it would be difficult to enumerate the amount of exploratory studies that have relied on these technologies or to find an environment whose microbiota and/or microbiome has not yet been at least preliminarily characterized. Indeed, many have begun reexamining classic paradigms of sterility surrounding various anatomical sites of the human body. Researchers have been investigating environments of progressively lower microbial biomass to the point where these investigations have reached the limits of detection for these molecular surveys [98, 103, 104, 119]. Numerous studies have now been published that fail to address and/or account for the limitations of sequencing technologies and this has shaped the various controversies laid out in this document, especially with respect to the existence of a placental and/or bladder microbiome.

In the context of the placenta, by pushing molecular techniques to their limits, premature conclusions have likely been drawn regarding the existence of a placental microbiome [44, 84, 95, 110-113]. If widely accepted, these conclusions have the potential to influence how clinicians perceive the microbiology of the placenta and the intra-amniotic environment and to alter current understanding of microbial invasion of the amniotic cavity and its effect on pregnancy outcomes [129, 136, 140, 143, 217, 303-307]. Our investigations of low microbial biomass sites address and account for the limitations of current molecular technologies through the inclusion of multiple microbiologic methodologies and extensive background technical controls.

We first investigated the human placenta and found no evidence to support the existence of a placental microbiome. We then extended our investigations to a mouse model. The value of a mouse model is that more than just the placenta can be sampled and interrogated for the existence of a microbiome – samples can be collected from the placenta, the fetus, and sites of the mother that are both typically colonized or presumed sterile. Again, we found no evidence of a placental microbiome, and further we found no evidence for *in utero* colonization of the fetus. Our results from our human and mouse model studies are largely congruent with the results of other recent studies [128, 167, 213, 225]: when appropriate measures are taken to address the caveats associated with investigations of low microbial biomass sites and the limitations of current technologies, there is no consistent evidence for a placental microbiome or *in utero* colonization of the fetus.

Alternatively, our investigations of the potential existence of a bladder microbiome, which also addressed and accounted for the limitations of current molecular technologies through the inclusion of multiple microbiologic methodologies and extensive background technical controls, revealed that the bladder is not likely a sterile organ. By following a similar study design to our human placental work [150], we successfully cultured bacteria from the bladders of pregnant women and detected molecular signals in the bladders through qPCR and 16S rRNA gene surveys that exceeded those of background technical controls. Additionally, by accounting for the possible influence of vulvovaginal contamination on the microbial profiles of urine samples, we were able to identify low abundance bacteria that are likely members of a bladder microbiota during pregnancy. *Ureaplasma* (detected through both culture and molecular surveys) and multiple Gram-positive anaerobic cocci (detected in molecular surveys, e.g *Finegoldia* and *Anaerococcus*), despite being present in low abundances, were identified as distinct features of

catheterized urine over clean catch urine and vaginal swabs. Despite the negative impacts *Ureaplasma* can have on pregnancy outcome in some women, it is regularly detected in the bladder and vagina of women that are asymptomatic [48, 248, 275], suggesting that while under some circumstances it may cause disease, it may also be a common commensal member of the microbial communities of these body sites. Although prior studies characterizing a bladder microbiome have not emphasized *Ureaplasma*, it is present at only low relative abundances in molecular surveys and may therefore be overlooked and it requires specific growth media which are typically not included in cultivation efforts of urine [115, 118]. Our research presented here provides a unique opportunity for future studies investigating *Ureaplasma* in the context of the bladder microbiome, as well elucidating potential genetic differences that may contribute to its being a commensal or a pathogen. The Gram-positive anaerobic cocci detected in our study are often detected in the bladder microbiome studies of others [48, 115, 118, 227, 275]. However, they are infrequently recovered in culture [118]. This is likely due to their fastidious growth requirements and long incubation periods [248, 274, 294]. These Gram-positive anaerobic cocci also present a unique opportunity for future investigations of their potential contributions to a bladder microbiome. Ultimately, efforts to identify and characterize the specific bacteria that are most closely associated with the uro-epithelium, such as intracellular or biofilm-producing bacteria, are necessary for developing a complete understanding of the bladder microenvironment and its potential role in human health and disease.

It is important to appreciate contemporary sequencing technologies for their facilitating our ability to characterize the microbial communities in all environments, including the human body. However, when seeking to overturn paradigms of sterility, we must employ multiple microbiologic methodologies and address and account for background DNA contamination. The

burden of proof for overturning longstanding paradigms of sterility is high; there must be clear and consistent evidence for the existence of a microbiome across multiple microbiologic methodologies and the signal of this microbiome must exceed that of controls and rise above the limits of detection of the different methodologies employed.

REFERENCES

1. Williams, D.W., *The Germ-Theory.* British medical journal, 1871. **1**(536): p. 368-368.
2. Lippi, D., Gotuzzo, E., and Caini, S., *Cholera.* Microbiol Spectr, 2016. **4**(4).
3. Pitt, D. and Aubin, J.-M., *Joseph Lister: father of modern surgery.* Canadian journal of surgery. Journal canadien de chirurgie, 2012. **55**(5): p. E8-E9.
4. Sanders, E.R., *Aseptic laboratory techniques: plating methods.* J Vis Exp, 2012(63): p. e3064.
5. Link, T., *Guideline Implementation: Sterile Technique.* Aorn j, 2019. **110**(4): p. 415-425.
6. Hart, S., *Using an aseptic technique to reduce the risk of infection.* Nursing Standard (through 2013), 2007. **21**(47): p. 43-8.
7. *Overhauling Aseptic Technique.* British Medical Journal, 1941. **1**(4180): p. 246-247.
8. Pennington, T.H., *Listerism, its decline and its persistence: the introduction of aseptic surgical techniques in three British teaching hospitals, 1890-99.* Medical history, 1995. **39**(1): p. 35-60.
9. Bykowski, T. and Stevenson, B., *Aseptic Technique.* Curr Protoc Microbiol, 2020. **56**(1): p. e98.
10. Coté, R.J., *Aseptic technique for cell culture.* Curr Protoc Cell Biol, 2001. **Chapter 1**: p. Unit 1.3.
11. Copley, J., *All at sea.* Nature, 2002. **415**(6872): p. 572-574.
12. *Microbiology by numbers.* Nature Reviews Microbiology, 2011. **9**(9): p. 628-628.
13. Louca, S., et al., *A census-based estimate of Earth's bacterial and archaeal diversity.* PLoS Biol, 2019. **17**(2): p. e3000106.

14. Prosser, J.I., et al., *The role of ecological theory in microbial ecology.* Nat Rev Microbiol, 2007. **5**(5): p. 384-92.

15. Troussellier, M. and Legendre, P., *A functional evenness index for microbial ecology.* Microb Ecol, 1981. **7**(4): p. 283-96.

16. Marchesi, J.R. and Ravel, J., *The vocabulary of microbiome research: a proposal.* Microbiome, 2015. **3**(1): p. 31.

17. Jannasch, H.W. and Mottl, M.J., *Geomicrobiology of Deep-Sea Hydrothermal Vents.* Science, 1985. **229**(4715): p. 717-725.

18. Mason, O.U., et al., *First Investigation of the Microbiology of the Deepest Layer of Ocean Crust.* PLoS One, 2010. **5**(11): p. e15399.

19. Kuang, J.-L., et al., *Contemporary environmental variation determines microbial diversity patterns in acid mine drainage.* The ISME Journal, 2013. **7**(5): p. 1038-1050.

20. Suzuki, S., et al., *Physiological and genomic features of highly alkaliphilic hydrogen-utilizing Betaproteobacteria from a continental serpentinizing site.* Nature Communications, 2014. **5**(1): p. 3900.

21. Sender, R., Fuchs, S., and Milo, R., *Revised Estimates for the Number of Human and Bacteria Cells in the Body.* PLoS Biol, 2016. **14**(8): p. e1002533.

22. Roser, M., Ritchie, H., and Ortiz-Ospina, E. *World Population Growth*. 2013; Available from: https://ourworldindata.org/world-population-growth.

23. Qin, J., et al., *A human gut microbial gene catalogue established by metagenomic sequencing.* Nature, 2010. **464**.

24. Snyder, W., et al., *Introduction.* Annals of the ICRP/ICRP Publication, 1975. **OS_23**(1): p. i-xix.

25. Sender, R., Fuchs, S., and Milo, R., *Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans.* Cell, 2016. **164**(3): p. 337-340.

26. Laupland, K.B. and Valiquette, L., *The changing culture of the microbiology laboratory.* Can J Infect Dis Med Microbiol, 2013. **24**(3): p. 125-8.

27. Didelot, X., et al., *Transforming clinical microbiology with bacterial genome sequencing.* Nat Rev Genet, 2012. **13**(9): p. 601-12.

28. Hamady, M. and Knight, R., *Microbial community profiling for human microbiome projects: Tools, techniques, and challenges.* Genome Res, 2009. **19**(7): p. 1141-52.

29. Hiergeist, A., et al., *Analyses of Intestinal Microbiota: Culture versus Sequencing.* ILAR J, 2015. **56**(2): p. 228-40.

30. Human Microbiome Project, C., *A framework for human microbiome research.* Nature, 2012. **486**(7402): p. 215-21.

31. Knight, R., et al., *The Microbiome and Human Biology.* Annu Rev Genomics Hum Genet, 2017.

32. Kell, D.B., et al., *Viability and activity in readily culturable bacteria: a review and discussion of the practical issues.* Antonie Van Leeuwenhoek, 1998. **73**(2): p. 169-87.

33. Greub, G., *Culturomics: a new approach to study the human microbiome.* Clin Microbiol Infect, 2012. **18**(12): p. 1157-9.

34. Browne, H.P., et al., *Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation.* Nature, 2016. **533**(7604): p. 543-6.

35. Stewart, E.J., *Growing unculturable bacteria.* J Bacteriol, 2012. **194**(16): p. 4151-60.

36. Vartoukian, S.R., Palmer, R.M., and Wade, W.G., *Strategies for culture of 'unculturable' bacteria.* FEMS Microbiol Lett, 2010. **309**(1): p. 1-7.

37. Spiegelman, D., Whissell, G., and Greer, C.W., *A survey of the methods for the characterization of microbial consortia and communities.* Can J Microbiol, 2005. **51**(5): p. 355-86.

38. Rhoads, D.D., et al., *Comparison of culture and molecular identification of bacteria in chronic wounds.* Int J Mol Sci, 2012. **13**(3): p. 2535-50.

39. Rhoads, D.D., et al., *Clinical identification of bacteria in human chronic wound infections: culturing vs. 16S ribosomal DNA sequencing.* BMC Infect Dis, 2012. **12**: p. 321.

40. Larsen, M.K., et al., *Use of cultivation-dependent and -independent techniques to assess contamination of central venous catheters: a pilot study.* BMC Clin Pathol, 2008. **8**: p. 10.

41. Schloss, P.D., et al., *Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities.* Appl Environ Microbiol, 2009. **75**(23): p. 7537-41.

42. NIH HMP Working Group, et al., *The NIH Human Microbiome Project.* Genome Res, 2009. **19**(12): p. 2317-23.

43. Ding, T. and Schloss, P.D., *Dynamics and associations of microbial community types across the human body.* Nature, 2014. **509**(7500): p. 357-60.

44. Aagaard, K., et al., *The placenta harbors a unique microbiome.* Science Translational Medicine, 2014. **6**(237): p. 237ra65.

45. Antony, K.M., et al., *The preterm placental microbiome varies in association with excess maternal gestational weight gain.* Am J Obstet Gynecol, 2015. **212**(5): p. 653 e1-16.

46. Arrieta, M.C., et al., *The intestinal microbiome in early life: health and disease.* Front Immunol, 2014. **5**: p. 427.

47. Chu, D.M., et al., *Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery.* Nat Med, 2017. **23**(3): p. 314-326.

48. Pearce, M.M., et al., *The female urinary microbiome: a comparison of women with and without urgency urinary incontinence.* MBio, 2014. **5**(4): p. e01283-14.

49. Prince, A.L., et al., *The placental membrane microbiome is altered among subjects with spontaneous preterm birth with and without chorioamnionitis.* Am J Obstet Gynecol, 2016. **214**(5): p. 627.e1-627.e16.

50. Romero, R., 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**: p. 18.

51. Romero, R., 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): p. 4.

52. Turnbaugh, P.J., et al., *A core gut microbiome in obese and lean twins.* Nature, 2009. **457**(7228): p. 480-4.

53. Turnbaugh, P.J., et al., *An obesity-associated gut microbiome with increased capacity for energy harvest.* Nature, 2006. **444**(7122): p. 1027-31.

54. Kassam, Z., et al., *Fecal microbiota transplantation for Clostridium difficile infection: systematic review and meta-analysis.* Am J Gastroenterol, 2013. **108**.

55. Mallozzi, M., Viswanathan, V.K., and Vedantam, G., *Spore-forming Bacilli and Clostridia in human disease.* Future Microbiology, 2010. **5**(7): p. 1109-1123.

56. Nood, E., et al., *Duodenal infusion of donor feces for recurrent Clostridium difficile.* N Engl J Med, 2013. **368**.

57. Wu, X., et al., *Molecular Characterisation of the Faecal Microbiota in Patients with Type II Diabetes.* Current Microbiology, 2010. **61**(1): p. 69-78.

58. Larsen, N., et al., *Gut Microbiota in Human Adults with Type 2 Diabetes Differs from Non-Diabetic Adults.* PLoS One, 2010. **5**(2): p. e9085.

59. Furet, J.-P., et al., *Differential Adaptation of Human Gut Microbiota to Bariatric Surgery–Induced Weight Loss.* Links With Metabolic and Low-Grade Inflammation Markers, 2010. **59**(12): p. 3049-3057.

60. Dashper, S.G., et al., *Temporal development of the oral microbiome and prediction of early childhood caries.* Scientific Reports, 2019. **9**(1): p. 19732.

61. Baker, J.L. and Edlund, A., *Exploiting the Oral Microbiome to Prevent Tooth Decay: Has Evolution Already Provided the Best Tools?* Frontiers in microbiology, 2019. **9**: p. 3323-3323.

62. Hurley, E., et al., *Comparison of the salivary and dentinal microbiome of children with severe-early childhood caries to the salivary microbiome of caries-free children.* BMC Oral Health, 2019. **19**(1): p. 13.

63. Brubaker, L., et al., *Urinary bacteria in adult women with urgency urinary incontinence.* Int Urogynecol J, 2014. **25**(9): p. 1179-84.

64. Brubaker, L. and Wolfe, A.J., *The female urinary microbiota, urinary health and common urinary disorders.* Ann Transl Med, 2017. **5**(2): p. 34.

65. Liu, F., et al., *Alterations in the Urinary Microbiota Are Associated With Cesarean Delivery.* Frontiers in Microbiology, 2018. **9**(2193).

66. Ollberding, N.J., et al., *Urinary Microbiota Associated with Preterm Birth: Results from the Conditions Affecting Neurocognitive Development and Learning in Early Childhood (CANDLE) Study.* PLoS One, 2016. **11**(9): p. e0162302.

67. Pearce, M.M., et al., *The female urinary microbiome in urgency urinary incontinence.* Am J Obstet Gynecol, 2015. **213**(3): p. 347 e1-11.

68. Thomas-White, K.J., et al., *Urinary microbes and postoperative urinary tract infection risk in urogynecologic surgical patients.* International urogynecology journal, 2018. **29**(12): p. 1797-1805.

69. Whiteside, S.A., et al., *The microbiome of the urinary tract--a role beyond infection.* Nat Rev Urol, 2015. **12**(2): p. 81-90.

70. Dickson, R.P., et al., *Changes in the Lung Microbiome following Lung Transplantation Include the Emergence of Two Distinct Pseudomonas Species with Distinct Clinical Associations.* Plos One, 2014. **9**(5): p. 15.

71. Chehoud, C., et al., *Fungal signature in the gut microbiota of pediatric patients with inflammatory bowel disease.* Inflamm Bowel Dis, 2015. **21**.

72. Cho, I. and Blaser, M.J., *The human microbiome: at the interface of health and disease.* Nat Rev Genet, 2012. **13**(4): p. 260-70.

73. Frank, D.N., et al., *Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases.* Proc Natl Acad Sci U S A, 2007. **104**(34): p. 13780-5.

74. Lee, D., et al., *Diet in the pathogenesis and treatment of inflammatory bowel diseases.* Gastroenterology, 2015. **148**.

75. Michail, S., et al., *Alterations in the gut microbiome of children with severe ulcerative colitis.* Inflamm Bowel Dis, 2012. **18**(10): p. 1799-808.

76. Round, J.L. and Mazmanian, S.K., *The gut microbiota shapes intestinal immune responses during health and disease.* Nat Rev Immunol, 2009. **9**(5): p. 313-23.

77. Shaw, K.A., et al., *Dysbiosis, inflammation, and response to treatment: a longitudinal study of pediatric subjects with newly diagnosed inflammatory bowel disease.* Genome Med, 2016. **8**.

78. Arrieta, M.C., et al., *Early infancy microbial and metabolic alterations affect risk of childhood asthma.* Sci Transl Med, 2015. **7**(307): p. 307ra152.

79. Hilty, M., et al., *Disordered microbial communities in asthmatic airways.* PLoS One, 2010. **5**(1): p. e8578.

80. Huang, Y.J. and Boushey, H.A., *The microbiome in asthma.* J Allergy Clin Immunol, 2015. **135**(1): p. 25-30.

81. Arumugam, M., et al., *Enterotypes of the human gut microbiome.* Nature, 2011. **473**.

82. Bäckhed, F., et al., *Dynamics and stabilization of the human gut microbiome during the first year of life.* Cell Host Microbe, 2015. **17**.

83. Borghi, E., et al., *Antenatal Microbial Colonization of Mammalian Gut.* Reproductive Sciences, 2019. **26**(8): p. 1045-1053.

84. Collado, M.C., et al., *Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid.* Sci Rep, 2016. **6**.

85. Dave, M., et al., *The human gut microbiome: current knowledge, challenges, and future directions.* Transl Res, 2012. **160**.

86. David, L.A., et al., *Diet rapidly and reproducibly alters the human gut microbiome.* Nature, 2014. **505**.

87. Debelius, J.W., et al., *Turning participatory microbiome research into usable data: lessons from the american gut project.* J Microbiol Biol Educ, 2016. **17**.

88. Dethlefsen, L., et al., *The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing.* PLoS Biol, 2008. **6**.

89. Dogra, S., et al., *Rate of establishing the gut microbiota in infancy has consequences for future health.* Gut Microbes, 2015. **6**.

90. Dollive, S., et al., *Fungi of the murine gut: episodic variation and proliferation during antibiotic treatment.* PLoS One, 2013. **8**.

91. Dore, J. and Blottiere, H., *The influence of diet on the gut microbiota and its consequences for health.* Curr Opin Biotechnol, 2015. **32**.

92. Fallucca, F., et al., *Influence of diet on gut microbiota, inflammation and type 2 diabetes mellitus. First experience with macrobiotic Ma-Pi 2 diet.* Diabetes Metab Res Rev, 2014. **30**.

93. Forslund, K., et al., *Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota.* Nature, 2015. **528**.

94. Gill, S.R., et al., *Metagenomic analysis of the human distal gut microbiome.* Science, 2006. **312**.

95. Gomez-Arango, L.F., et al., *Contributions of the maternal oral and gut microbiome to placental microbial colonization in overweight and obese pregnant women.* Scientific Reports, 2017. **7**: p. 10.

96. Ley, R.E., et al., *Microbial ecology: human gut microbes associated with obesity.* Nature, 2006. **444**(7122): p. 1022-3.

97. Caruso, V., et al., *Performance of Microbiome Sequence Inference Methods in Environments with Varying Biomass.* mSystems, 2019. **4**(1): p. e00163-18.

98. Kim, D., et al., *Optimizing methods and dodging pitfalls in microbiome research.* Microbiome, 2017. **5**(1): p. 52.

99. Salter, S.J., et al., *Reagent and laboratory contamination can critically impact sequence-based microbiome analyses.* BMC Biol, 2014. **12**.

100. de Goffau, M.C., et al., *Recognizing the reagent microbiome.* Nat Microbiol, 2018. **3**(8): p. 851-853.

101. Jervis-Bardy, J., et al., *Deriving accurate microbiota profiles from human samples with low bacterial content through post-sequencing processing of Illumina MiSeq data.* Microbiome, 2015. **3**.

102. Laurence, M., Hatzis, C., and Brash, D.E., *Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes.* PLoS One, 2014. **9**.

103. O'Callaghan, J., et al., *Re-assessing microbiomes in the low-biomass reproductive niche.* BJOG: An International Journal of Obstetrics & Gynaecology, 2020. **127**(2): p. 147-158.

104. Karstens, L., et al., *Controlling for Contaminants in Low-Biomass 16S rRNA Gene Sequencing Experiments.* mSystems, 2019. **4**(4): p. e00290-19.

105. Emery, D.C., et al., *16S rRNA Next Generation Sequencing Analysis Shows Bacteria in Alzheimer's Post-Mortem Brain.* Frontiers in Aging Neuroscience, 2017. **9**(195).

106. Païssé, S., et al., *Comprehensive description of blood microbiome from healthy donors assessed by 16S targeted metagenomic sequencing.* Transfusion, 2016. **56**(5): p. 1138-1147.

107. Potgieter, M., et al., *The dormant blood microbiome in chronic, inflammatory diseases.* FEMS Microbiology Reviews, 2015. **39**(4): p. 567-591.

108. Moreno, I., et al., *Evidence that the endometrial microbiota has an effect on implantation success or failure.* Am J Obstet Gynecol, 2016. **215**(6): p. 684-703.

109. Winters, A.D., et al., *Does the endometrial cavity have a molecular microbial signature?* Sci Rep, 2019. **9**(1): p. 9905.

110. Doyle, R.M., et al., *Term and preterm labour are associated with distinct microbial community structures in placental membranes which are independent of mode of delivery.* Placenta, 2014. **35**.

111. Zheng, J., et al., *The placental microbiome varies in association with low birth weight in full-term neonates.* Nutrients, 2015. **7**.

112. Bassols, J., et al., *Gestational diabetes is associated with changes in placental microbiota and microbiome.* Pediatr Res, 2016. **80**(6): p. 777-784.

113. Parnell, L.A., et al., *Microbial communities in placentas from term normal pregnancy exhibit spatially variable profiles.* Sci Rep, 2017. **7**(1): p. 11200.

114. Brubaker, L. and Wolfe, A.J., *The new world of the urinary microbiota in women.* Am J Obstet Gynecol, 2015. **213**(5): p. 644-9.

115. Hilt, E.E., et al., *Urine is not sterile: use of enhanced urine culture techniques to detect resident bacterial flora in the adult female bladder.* J Clin Microbiol, 2014. **52**(3): p. 871-6.

116. Siddiqui, H., et al., *Assessing diversity of the female urine microbiota by high throughput sequencing of 16S rDNA amplicons.* BMC Microbiol, 2011. **11**: p. 244.

117. Wolfe, A.J. and Brubaker, L., *"Sterile Urine" and the Presence of Bacteria.* Eur Urol, 2015. **68**(2): p. 173-4.

118. Wolfe, A.J., et al., *Evidence of uncultivated bacteria in the adult female bladder.* J Clin Microbiol, 2012. **50**(4): p. 1376-83.

119. Glassing, A., et al., *Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples.* Gut Pathog, 2016. **8**.

120. Davis, N.M., et al., *Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data.* Microbiome, 2018. **6**(1): p. 226.

121. Knights, D., et al., *Bayesian community-wide culture-independent microbial source tracking.* Nat Methods, 2011. **8**.

122. Segata, N., et al., *Metagenomic biomarker discovery and explanation.* Genome Biology, 2011. **12**(6): p. 18.

123. Frank, J.A., et al., *Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes.* Applied and Environmental Microbiology, 2008. **74**(8): p. 2461-2470.

124. Klindworth, A., et al., *Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies.* Nucleic Acids Res, 2013. **41**(1): p. e1.

125. Wang, Y. and Qian, P.-Y., *Conservative Fragments in Bacterial 16S rRNA Genes and Primer Design for 16S Ribosomal DNA Amplicons in Metagenomic Studies.* PLoS One, 2009. **4**(10): p. e7401.

126. Volpi, E.V. and Bridger, J.M., *FISH glossary: an overview of the fluorescence in situ hybridization technique.* BioTechniques, 2008. **45**(4): p. 385-409.

127. Levsky, J.M. and Singer, R.H., *Fluorescence in situ hybridization: past, present and future.* Journal of Cell Science, 2003. **116**(14): p. 2833-2838.

128. Lauder, A.P., et al., *Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota.* Microbiome, 2016. **4**.

129. Blanc, W.A., *Amniotic infection syndrome; pathogenesis, morphology, and significance in circumnatal mortality.* Clin Obstet Gynecol, 1959. **2**: p. 705-34.

130. Fernandez-Feo, M., et al., *The cultivable human oral gluten-degrading microbiome and its potential implications in coeliac disease and gluten sensitivity.* Clin Microbiol Infect, 2013. **19**.

131. Garrett, W.S., *Cancer and the microbiota.* Science, 2015. **348**(6230): p. 80-6.

132. Lewis, J.D., et al., *Inflammation, antibiotics, and diet as environmental stressors of the gut microbiome in pediatric crohn's disease.* Cell Host Microbe, 2015. **18**.

133. Thomas-White, K., et al., *Culturing of female bladder bacteria reveals an interconnected urogenital microbiota.* Nat Commun, 2018. **9**(1): p. 1557.

134. Gilstrap, L.C., 3rd, Cunningham, F.G., and Whalley, P.J., *Acute pyelonephritis in pregnancy: an anterospective study.* Obstet Gynecol, 1981. **57**(4): p. 409-13.

135. Eschenbach, D.A., et al., *Bacterial vaginosis during pregnancy. An association with prematurity and postpartum complications.* Scand J Urol Nephrol Suppl, 1984. **86**: p. 213-22.

136. Gravett, M.G., et al., *Preterm labor associated with subclinical amniotic fluid infection and with bacterial vaginosis.* Obstet Gynecol, 1986. **67**(2): p. 229-37.

137. Patterson, T.F. and Andriole, V.T., *Bacteriuria in pregnancy.* Infect Dis Clin North Am, 1987. **1**(4): p. 807-22.

138. Romero, R. and Mazor, M., *Infection and preterm labor.* Clin Obstet Gynecol, 1988. **31**(3): p. 553-84.

139. Hillier, S.L., et al., *A case-control study of chorioamnionic infection and histologic chorioamnionitis in prematurity.* N Engl J Med, 1988. **319**(15): p. 972-8.

140. Romero, R., et al., *Infection and labor. VI. Prevalence, microbiology, and clinical significance of intraamniotic infection in twin gestations with preterm labor.* Am J Obstet Gynecol, 1990. **163**(3): p. 757-61.

141. Hillier, S.L., et al., *Microbiologic causes and neonatal outcomes associated with chorioamnion infection.* Am J Obstet Gynecol, 1991. **165**(4 Pt 1): p. 955-61.

142. Andrews, W.W., Goldenberg, R.L., and Hauth, J.C., *Preterm labor: emerging role of genital tract infections.* Infectious Agents and Disease, 1995. **4**(4): p. 196-211.

143. Horowitz, S., et al., *Infection of the amniotic cavity with Ureaplasma urealyticum in the midtrimester of pregnancy.* J Reprod Med, 1995. **40**.

144. Goldenberg, R.L., Hauth, J.C., and Andrews, W.W., *Intrauterine infection and preterm delivery.* N Engl J Med, 2000. **342**(20): p. 1500-7.

145. Yoon, B.H., et al., *Clinical implications of detection of Ureaplasma urealyticum in the amniotic cavity with the polymerase chain reaction.* Am J Obstet Gynecol, 2000. **183**(5): p. 1130-7.

146. Romero, R., et al., *The role of infection in preterm labour and delivery.* Paediatr Perinat Epidemiol, 2001. **15 Suppl 2**: p. 41-56.

147. Falagas, M.E., Betsi, G.I., and Athanasiou, S., *Probiotics for the treatment of women with bacterial vaginosis.* Clin Microbiol Infect, 2007. **13**.

148. Rautava, S., et al., *Probiotics modulate host-microbe interaction in the placenta and fetal gut: a randomized, double-blind, placebo-controlled trial.* Neonatology, 2012. **102**.

149. Perez-Munoz, M.E., et al., *A critical assessment of the "sterile womb" and "in utero colonization" hypotheses: implications for research on the pioneer infant microbiome.* Microbiome, 2017. **5**: p. 19.

150. Theis, K.R., et al., *Does the human placenta delivered at term have a microbiota? Results of cultivation, quantitative real-time PCR, 16S rRNA gene sequencing, and metagenomics.* American Journal of Obstetrics & Gynecology, 2019. **220**(3): p. 267.e1-267.e39.

151. Schloss, P.D. and Handelsman, J., *Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness.* Applied and Environmental Microbiology, 2005. **71**.

152. Callahan, B.J., et al., *DADA2: High-resolution sample inference from Illumina amplicon data.* Nature Methods, 2016. **13**(7): p. 581-583.

153. Lewis, D.A., et al., *The human urinary microbiome; bacterial DNA in voided urine of asymptomatic adults.* Front Cell Infect Microbiol, 2013. **3**: p. 41.

154. Kogan, M.I., et al., *Human urine is not sterile - shift of paradigm.* Urol Int, 2015. **94**(4): p. 445-52.

155. Colgan, R. and Williams, M., *Diagnosis and treatment of acute uncomplicated cystitis.* Am Fam Physician, 2011. **84**(7): p. 771-6.

156. Gilbert, G.L., et al., *Bacteriuria due to ureaplasmas and other fastidious organisms during pregnancy: prevalence and significance.* Pediatr Infect Dis, 1986. **5**(6 Suppl): p. S239-43.

157. Pfau, A. and Sacks, T.G., *Effective prophylaxis for recurrent urinary tract infections during pregnancy.* Clin Infect Dis, 1992. **14**(4): p. 810-4.

158. Delzell, J.E., Jr. and Lefevre, M.L., *Urinary tract infections during pregnancy.* Am Fam Physician, 2000. **61**(3): p. 713-21.

159. Schnarr, J. and Smaill, F., *Asymptomatic bacteriuria and symptomatic urinary tract infections in pregnancy.* Eur J Clin Invest, 2008. **38 Suppl 2**: p. 50-7.

160. Farkash, E., et al., *Acute antepartum pyelonephritis in pregnancy: a critical analysis of risk factors and outcomes.* Eur J Obstet Gynecol Reprod Biol, 2012. **162**(1): p. 24-7.

161. Piatek, J., et al., *Urinary tract infection during pregnancy affects the level of leptin, ghrelin and insulin in maternal and placental blood.* Scand J Clin Lab Invest, 2014. **74**(2): p. 126-31.

162. Glaser, A.P. and Schaeffer, A.J., *Urinary Tract Infection and Bacteriuria in Pregnancy.* Urol Clin North Am, 2015. **42**(4): p. 547-60.

163. Matuszkiewicz-Rowinska, J., Malyszko, J., and Wieliczko, M., *Urinary tract infections in pregnancy: old and new unresolved diagnostic and therapeutic problems.* Arch Med Sci, 2015. **11**(1): p. 67-77.

164. Easter, S.R., et al., *Urinary tract infection during pregnancy, angiogenic factor profiles, and risk of preeclampsia.* Am J Obstet Gynecol, 2016. **214**(3): p. 387 e1-7.

165. Romero, R., et al., *Meta-analysis of the relationship between asymptomatic bacteriuria and preterm delivery/low birth weight.* Obstet Gynecol, 1989. **73**(4): p. 576-82.

166. Leuck, A.-M., et al., *Complications of Foley Catheters—Is Infection the Greatest Risk?* The Journal of Urology, 2012. **187**(5): p. 1662-1666.

167. Leiby, J.S., et al., *Lack of detection of a human placenta microbiome in samples from preterm and term deliveries.* Microbiome, 2018. **6**(1): p. 196.

168. Chen, C., et al., *The microbiota continuum along the female reproductive tract and its relation to uterine-related diseases.* Nat Commun, 2017. **8**(1): p. 875.

169. Jacobs, K.M., et al., *Microorganisms Identified in the Maternal Bladder: Discovery of the Maternal Bladder Microbiota.* AJP Rep, 2017. **7**(3): p. e188-e196.

170. Lamont, R.F., et al., *The vaginal microbiome: new information about genital tract flora using molecular based techniques.* BJOG, 2011. **118**(5): p. 533-49.

171. Ravel, J., et al., *Vaginal microbiome of reproductive-age women.* Proc Natl Acad Sci U S A, 2011. **108 Suppl 1**: p. 4680-7.

172. Gajer, P., et al., *Temporal dynamics of the human vaginal microbiota.* Sci Transl Med, 2012. **4**.

173. Elovitz, M.A., et al., *Cervicovaginal microbiota and local immune response modulate the risk of spontaneous preterm delivery.* Nat Commun, 2019. **10**(1): p. 1305.

174. Fettweis, J.M., et al., *The vaginal microbiome and preterm birth.* Nat Med, 2019. **25**(6): p. 1012-1021.

175. Seferovic, M.D., et al., *Visualization of microbes by 16S in situ hybridization in term and preterm placentas without intraamniotic infection.* American Journal of Obstetrics & Gynecology, 2019. **221**(2): p. 146.e1-146.e23.

176. Robbins, J.R. and Bakardjiev, A.I., *Pathogens and the placental fortress.* Curr Opin Microbiol, 2012. **15**.

177. DiGiulio, D.B., et al., *Prevalence and diversity of microbes in the amniotic fluid, the fetal inflammatory response, and pregnancy outcome in women with preterm pre-labor rupture of membranes.* Am J Reprod Immunol, 2010. **64**(1): p. 38-57.

178. Blanc, W.A., *Pathways of fetal and early neonatal infection. Viral placentitis, bacterial and fungal chorioamnionitis.* J Pediatr, 1961. **59**: p. 473-96.

179. Aquino, T.I., et al., *Subchorionic fibrin cultures for bacteriologic study of the placenta.* Am J Clin Pathol, 1984. **81**.

180. Svensson, L., Ingemarsson, I., and MÅrdh, P.-a., *Chorioamnionitis and the isolation of microorganisms from the placenta.* Obstetrics & Gynecology, 1986. **67**(3): p. 403-409.

181. Ovalle, A., et al., *Placental histopathology in premature rupture of membranes. Its relationship with microbiological findings, maternal, and neonatal outcome.* Rev Médica Chile, 1998. **126**.

182. Stout, M.J., et al., *Identification of intracellular bacteria in the basal plate of the human placenta in term and preterm gestations.* Am J Obstet Gynecol, 2013. **208**(3): p. 226 e1-7.

183. Prince, A.L., et al., *The perinatal microbiome and pregnancy: moving beyond the vaginal microbiome.* Cold Spring Harb Perspect Med, 2015. **5**(6).

184. Funkhouser, L.J. and Bordenstein, S.R., *Mom knows best: the universality of maternal microbial transmission.* PLoS Biol, 2013. **11**.

185. Pelzer, E., et al., *Review: Maternal health and the placental microbiome.* Placenta, 2017. **54**: p. 30-37.

186. Koleva, P.T., et al., *Microbial programming of health and disease starts during fetal life.* Birth Defects Res C Embryo Today, 2015. **105**.

187. Cao, B., et al., *Placental Microbiome and Its Role in Preterm Birth.* Neoreviews, 2014. **15**(12): p. e537-e545.

188. Zaura, E., et al., *Acquiring and maintaining a normal oral microbiome: current perspective.* Front Cell Infect Microbiol, 2014. **4**: p. 85.

189. Weiss, S., et al., *Tracking down the sources of experimental contamination in microbiome studies.* Genome Biol, 2014. **15**.

190. Caporaso, J.G., et al., *QIIME allows analysis of high-throughput community sequencing data.* Nat Methods, 2010. **7**.

191. Nearing, J.T., et al., *Denoising the Denoisers: an independent evaluation of microbiome sequence error-correction approaches.* PeerJ, 2018. **6**: p. e5364.

192. Callahan, B.J., McMurdie, P.J., and Holmes, S.P., *Exact sequence variants should replace operational taxonomic units in marker-gene data analysis.* Isme j, 2017. **11**(12): p. 2639-2643.

193. R Core Team, *R: A Language and Environment for Statistical Computing*. 2020, R Foundation for Statistical Computing: Vienna, Austria.

194. Meyer, F., et al., *The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes.* BMC Bioinformatics, 2008. **9**: p. 386.

195. Mason, O.U., et al., *Metagenomics reveals sediment microbial community response to Deepwater Horizon oil spill.* ISME J, 2014. **8**(7): p. 1464-1475.

196. Kopylova, E., Noe, L., and Touzet, H., *SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data.* Bioinformatics, 2012. **28**(24): p. 3211-3217.

197. Fu, L.M., et al., *CD-HIT: accelerated for clustering the next-generation sequencing data.* Bioinformatics, 2012. **28**(23): p. 3150-3152.

198. Kent, W.J., *BLAT - The BLAST-like alignment tool.* Genome Research, 2002. **12**(4): p. 656-664.

199. Rho, M.N., Tang, H.X., and Ye, Y.Z., *FragGeneScan: predicting genes in short and error-prone reads.* Nucleic Acids Research, 2010. **38**(20): p. 12.

200. Kanehisa, M., et al., *KEGG as a reference resource for gene and protein annotation.* Nucleic Acids Research, 2016. **44**(D1): p. D457-D462.

201. RStudio Team, *RStudio: Integrated Development for R.* 2020, "RStudio, PBC": Boston, MA.

202. Hammer, Ø., Harper, D.A.T., and Ryan, P.D., *PAST: Paleontological statistics software package for education and data analysis.* Palaeontologia Electronica, 2001. **4**(1): p. 1-9.

203. Dixon, P., *VEGAN, a package of R functions for community ecology.* Journal of Vegetation Science, 2003. **14**.

204. Broad Institute, *Morpheus*. 2020, Broad Institute.

205. Johnson, M., et al., *NCBI BLAST: a better web interface.* Nucleic acids research, 2008. **36**(suppl_2): p. W5-W9.

206. Zheng, J., et al., *The Placental Microbiota Is Altered among Subjects with Gestational Diabetes Mellitus: A Pilot Study.* Frontiers in physiology, 2017. **8**: p. 675-675.

207. Leon, L.J., et al., *Enrichment of Clinically Relevant Organisms in Spontaneous Preterm-Delivered Placentas and Reagent Contamination across All Clinical Groups in a Large Pregnancy Cohort in the United Kingdom.* Appl Environ Microbiol, 2018. **84**(14).

208. Kageyama, A. and Benno, Y., *Coprobacillus catenaformis Gen. Nov., Sp. Nov., a New Genus and Species Isolated from Human Feces.* Microbiology and Immunology, 2000. **44**(1): p. 23-28.

209. Cole, J.R., et al., *Ribosomal Database Project: data and tools for high throughput rRNA analysis.* Nucleic Acids Res, 2014. **42**(Database issue): p. D633-42.

210. Doyle, R.M., et al., *Bacterial communities found in placental tissues are associated with severe chorioamnionitis and adverse birth outcomes.* PLoS One, 2017. **12**(7): p. e0180167.

211. Marsh, R.L., et al., *How low can we go? The implications of low bacterial load in respiratory microbiota studies.* Pneumonia, 2018. **10**(1): p. 7.

212. Theis, K.R., et al., *No Consistent Evidence for Microbiota in Murine Placental and Fetal Tissues.* mSphere, 2020. **5**(1): p. e00933-19.

213. Younge, N., et al., *Fetal exposure to the maternal microbiota in humans and mice.* JCI Insight, 2019. **4**(19).

214. Martinez, K.A., II, et al., *Bacterial DNA is present in the fetal intestine and overlaps with that in the placenta in mice.* PLoS One, 2018. **13**(5): p. e0197439.

215. Amarasekara, R., et al., *Microbiome of the placenta in pre-eclampsia supports the role of bacteria in the multifactorial cause of pre-eclampsia.* J Obstet Gynaecol Res, 2015. **41**.

216. Egli, G.E. and Newton, M., *The transport of carbon particles in the human female reproductive tract.* Fertil Steril, 1961. **12**: p. 151-5.

217. Romero, R., et al., *Microbial invasion of the amniotic cavity during term labor. Prevalence and clinical significance.* J Reprod Med, 1993. **38**(7): p. 543-8.

218. Seong, H.S., et al., *The frequency of microbial invasion of the amniotic cavity and histologic chorioamnionitis in women at term with intact membranes in the presence or absence of labor.* Am J Obstet Gynecol, 2008. **199**(4): p. 375.e1-5.

219. Kozich, J.J., et al., *Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform.* Appl Environ Microbiol, 2013. **79**(17): p. 5112-20.

220. Gallup, J., *qPCR inhibition and amplification of difficult templates*, in *PCR troubleshooting and optimization: the essential guide*, S. Kennedy and N. Oswald, Editors. 2001, Caister Academic Press: Norfolk, UK. p. 1-59.

221. Pruesse, E., et al., *SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB.* Nucleic acids research, 2007. **35**(21): p. 7188-7196.

222. Quast, C., et al., *The SILVA ribosomal RNA gene database project: improved data processing and web-based tools.* Nucleic Acids Res, 2013. **41**(Database issue): p. D590-6.

223. Rognes, T., et al., *VSEARCH: a versatile open source tool for metagenomics.* PeerJ, 2016. **4**: p. e2584.

224. Lagkouvardos, I., et al., *The Mouse Intestinal Bacterial Collection (miBC) provides host-specific insight into cultured diversity and functional potential of the gut microbiota.* Nature Microbiology, 2016. **1**(10): p. 16131.

225. Kuperman, A., et al., *Deep microbial analysis of multiple placentas shows no evidence for a placental microbiome.* BJOG: An International Journal of Obstetrics & Gynaecology, 2020. **127**(2): p. 159-169.

226. Lakeman, M.M. and Roovers, J.P., *Urinary tract infections in women with urogynaecological symptoms.* Curr Opin Infect Dis, 2016. **29**(1): p. 92-7.

227. Thomas-White, K., et al., *The bladder is not sterile: History and current discoveries on the urinary microbiome.* Curr Bladder Dysfunct Rep, 2016. **11**(1): p. 18-24.

228. Basu, S., et al., *Evolution of bacterial and fungal growth media.* Bioinformation, 2015. **11**(4): p. 182-4.

229. Stamm, W.E. and Hooton, T.M., *Management of urinary tract infections in adults.* N Engl J Med, 1993. **329**(18): p. 1328-34.

230. Chin, T.L., et al., *Antimicrobial resistance surveillance in urinary tract infections in primary care.* J Antimicrob Chemother, 2016. **71**(10): p. 2723-8.

231. Le, J., et al., *Urinary tract infections during pregnancy.* Ann Pharmacother, 2004. **38**(10): p. 1692-701.

232. Souza, R.B., Trevisol, D.J., and Schuelter-Trevisol, F., *Bacterial sensitivity to fosfomycin in pregnant women with urinary infection.* Braz J Infect Dis, 2015. **19**(3): p. 319-23.

233. Szweda, H. and Jozwik, M., *Urinary tract infections during pregnancy - an updated overview.* Dev Period Med, 2016. **20**(4): p. 263-272.

234. Chaemsaithong, P., et al., *Soluble TRAIL in normal pregnancy and acute pyelonephritis: a potential explanation for the susceptibility of pregnant women to microbial products and infection.* J Matern Fetal Neonatal Med, 2013. **26**(16): p. 1568-75.

235. Romero, R., et al., *Inflammation in pregnancy: its roles in reproductive physiology, obstetrical complications, and fetal injury.* Nutr Rev, 2007. **65**(12 Pt 2): p. S194-202.

236. Gervasi, M.T., et al., *Maternal intravascular inflammation in preterm premature rupture of membranes.* J Matern Fetal Neonatal Med, 2002. **11**(3): p. 171-5.

237. Gervasi, M.T., et al., *Phenotypic and metabolic characteristics of monocytes and granulocytes in preeclampsia.* Am J Obstet Gynecol, 2001. **185**(4): p. 792-7.

238. Gervasi, M.T., et al., *Phenotypic and metabolic characteristics of maternal monocytes and granulocytes in preterm labor with intact membranes.* Am J Obstet Gynecol, 2001. **185**(5): p. 1124-9.

239. Soto, E., et al., *Fragment Bb: evidence for activation of the alternative pathway of the complement system in pregnant women with acute pyelonephritis.* J Matern Fetal Neonatal Med, 2010. **23**(10): p. 1085-90.

240. Chaiworapongsa, T., et al., *Acute pyelonephritis during pregnancy changes the balance of angiogenic and anti-angiogenic factors in maternal plasma.* J Matern Fetal Neonatal Med, 2010. **23**(2): p. 167-78.

241. Gotsch, F., et al., *Maternal serum concentrations of the chemokine CXCL10/IP-10 are elevated in acute pyelonephritis during pregnancy.* J Matern Fetal Neonatal Med, 2007. **20**(10): p. 735-44.

242. Soto, E., et al., *Increased concentration of the complement split product C5a in acute pyelonephritis during pregnancy.* J Matern Fetal Neonatal Med, 2005. **17**(4): p. 247-52.

243. Kusanovic, J.P., et al., *Maternal serum soluble CD30 is increased in pregnancies complicated with acute pyelonephritis.* J Matern Fetal Neonatal Med, 2007. **20**(11): p. 803-11.

244. Widmer, M., et al., *Duration of treatment for asymptomatic bacteriuria during pregnancy.* Cochrane Database Syst Rev, 2015(11): p. CD000491.

245. Sweet, R.L., *Bacteriuria and pyelonephritis during pregnancy.* Semin Perinatol, 1977. **1**(1): p. 25-40.

246. Patterson, T.F. and Andriole, V.T., *Detection, significance, and therapy of bacteriuria in pregnancy. Update in the managed health care era.* Infect Dis Clin North Am, 1997. **11**(3): p. 593-608.

247. Nicolle, L.E., et al., *Infectious Diseases Society of America guidelines for the diagnosis and treatment of asymptomatic bacteriuria in adults.* Clin Infect Dis, 2005. **40**(5): p. 643-54.

248. McDowall, D.R.M., et al., *Anaerobic and Other Fastidious Microorganisms in Asymptomatic Bacteriuria in Pregnant Women.* The Journal of Infectious Diseases, 1981. **144**(2): p. 114-122.

249. Smaill, F.M. and Vazquez, J.C., *Antibiotics for asymptomatic bacteriuria in pregnancy.* Cochrane Database Syst Rev, 2015(8): p. CD000490.

250. Chu, C.M. and Lowder, J.L., *Diagnosis and Treatment of Urinary Tract Infections Across Age Groups.* Am J Obstet Gynecol, 2018.

251. Redelinghuys, M.J., et al., *Comparison of the new Mycofast Revolution assay with a molecular assay for the detection of genital mycoplasmas from clinical specimens.* BMC Infect Dis, 2013. **13**: p. 453.

252. Singhal, N., et al., *MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis.* Front Microbiol, 2015. **6**: p. 791.

253. Caporaso, J.G., et al., *Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms.* ISME J, 2012. **6**(8): p. 1621-1624.

254. Edgar, R.C., et al., *UCHIME improves sensitivity and speed of chimera detection.* Bioinformatics, 2011. **27**(16): p. 2194-200.

255. Wang, Q., et al., *Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy.* Applied and environmental microbiology, 2007. **73**(16): p. 5261-5267.

256. Højsgaard, S., Halekoh, U., and Yan, J., *The R Package geepack for Generalized Estimating Equations.* 2005, 2005. **15**(2): p. 11.

257. Anderson, M.J., *A new method for non-parametric multivariate analysis of variance.* Austral Ecology, 2001. **26**(1): p. 32-46.

258. Quinn, G.P. and Keough, M.J., *Experimental Design and Data Analysis for Biologists*. 2002, Cambridge: Cambridge University Press.

259. Gotelli, N.J. and Ellison, A.M., *A Primer of Ecological Statistics*. 2004, Sunderland, MA: Sinauer Associates, Inc.

260. Foxman, B., et al., *Urinary tract infection: self-reported incidence and associated costs.* Ann Epidemiol, 2000. **10**(8): p. 509-15.

261. Fihn, S.D., *Clinical practice. Acute uncomplicated urinary tract infection in women.* N Engl J Med, 2003. **349**(3): p. 259-66.

262. American College of, O. and Gynecologists, *ACOG Practice Bulletin No. 91: Treatment of urinary tract infections in nonpregnant women.* Obstet Gynecol, 2008. **111**(3): p. 785-94.

263. Foxman, B., *Recurring urinary tract infection: incidence and risk factors.* Am J Public Health, 1990. **80**(3): p. 331-3.

264. Loh, K. and Sivalingam, N., *Urinary tract infections in pregnancy.* Malays Fam Physician, 2007. **2**(2): p. 54-7.

265. Kranz, J., et al., *The 2017 Update of the German Clinical Guideline on Epidemiology, Diagnostics, Therapy, Prevention, and Management of Uncomplicated Urinary Tract Infections in Adult Patients: Part 1.* Urol Int, 2018.

266. ACOG. Available from: https://www.acog.org/Patients/FAQs/Routine-Tests-During-Pregnancy#why.

267. *ACOG educational bulletin. Antimicrobial therapy for obstetric patients. Number 245, March 1998 (replaces no. 117, June 1988). American College of Obstetricians and Gynecologists.* Int J Gynaecol Obstet, 1998. **61**(3): p. 299-308.

268. Schneeberger, C., et al., *Contamination rates of three urine-sampling methods to assess bacteriuria in pregnant women.* Obstet Gynecol, 2013. **121**(2 Pt 1): p. 299-305.

269. Baerheim, A., Digranes, A., and Hunskaar, S., *Evaluation of urine sampling technique: bacterial contamination of samples from women students.* Br J Gen Pract, 1992. **42**(359): p. 241-3.

270. Lifshitz, E. and Kramer, L., *Outpatient urine culture: does collection technique matter?* Arch Intern Med, 2000. **160**(16): p. 2537-40.

271. Abdel-Aleem, H., et al., *Indwelling bladder catheterisation as part of intraoperative and postoperative care for caesarean section.* Cochrane Database Syst Rev, 2014(4): p. CD010322.

272. Acharya, S., et al., *Cesarean section without urethral catheterization: a randomized control trial.* Kathmandu Univ Med J (KUMJ), 2012. **10**(38): p. 18-22.

273. Nasr, A.M., et al., *Evaluation of the use vs nonuse of urinary catheterization during cesarean delivery: a prospective, multicenter, randomized controlled trial.* J Perinatol, 2009. **29**(6): p. 416-21.

274. Ezaki, T., Li, N., and Kawamura, Y., *The Anaerobic Gram-Positive Cocci*, in *The Prokaryotes: Volume 4: Bacteria: Firmicutes, Cyanobacteria*, M. Dworkin, et al., Editors. 2006, Springer US: New York, NY. p. 795-808.

275. Siddiqui, H., et al., *Alterations of microbiota in urine from women with interstitial cystitis.* BMC microbiology, 2012. **12**: p. 205-205.

276. Pagnier, I., et al., *Non-contiguous finished genome sequence and description of Fenollaria massiliensis gen. nov., sp. nov., a new genus of anaerobic bacterium.* Standards in genomic sciences, 2014. **9**(3): p. 704-717.

277. Boiten, K.E., Jean-Pierre, H., and Veloo, A.C.M., *Assessing the clinical relevance of Fenollaria massiliensis in human infections, using MALDI-TOF MS.* Anaerobe, 2018. **54**: p. 240-245.

278. Birch, D.F., Fairley, K.F., and Pavillard, R.E., *Unconventional bacteria in urinary tract disease: Ureaplasma urealyticum*. Kidney International, 1981. **19**(1): p. 58-64.

279. Schneeberger, C., Kazemier, B.M., and Geerlings, S.E., *Asymptomatic bacteriuria and urinary tract infections in special patient groups: women with diabetes mellitus and pregnant women.* Curr Opin Infect Dis, 2014. **27**(1): p. 108-14.

280. Cox, C., et al., *The common vaginal commensal bacterium Ureaplasma parvum is associated with chorioamnionitis in extreme preterm labor.* Journal of Maternal-Fetal & Neonatal Medicine, 2016. **29**(22): p. 3646-3651.

281. DiGiulio, D.B., *Diversity of microbes in amniotic fluid.* Semin Fetal Neonatal Med, 2012. **17**(1): p. 2-11.

282. Murtha, A.P. and Edwards, J.M., *The Role of Mycoplasma and Ureaplasma in Adverse Pregnancy Outcomes.* Obstetrics and Gynecology Clinics of North America, 2014. **41**(4): p. 615-+.

283. Romero, R., et al., *Sterile and microbial-associated intra-amniotic inflammation in preterm prelabor rupture of membranes.* J Matern Fetal Neonatal Med, 2015. **28**(12): p. 1394-409.

284. Blanchard, A., et al., *Detection of Ureaplasma urealyticum by polymerase chain reaction in the urogenital tract of adults, in amniotic fluid, and in the respiratory tract of newborns.* Clin Infect Dis, 1993. **17 Suppl 1**: p. S148-53.

285. Izraeli, S., et al., *Genital mycoplasmas in preterm infants: prevalence and clinical significance.* Eur J Pediatr, 1991. **150**(11): p. 804-7.

286. Ollikainen, J., et al., *Ureaplasma urealyticum infection associated with acute respiratory insufficiency and death in premature infants.* J Pediatr, 1993. **122**(5 Pt 1): p. 756-60.

287. Cassell, G.H., et al., *Association of Ureaplasma urealyticum infection of the lower respiratory tract with chronic lung disease and death in very-low-birth-weight infants.* Lancet, 1988. **2**(8605): p. 240-5.

288. Romero, R., et al., *The preterm parturition syndrome.* BJOG, 2006. **113 Suppl 3**: p. 17-42.

289. Xiao, L., et al., *Detection and characterization of human Ureaplasma species and serovars by real-time PCR.* J Clin Microbiol, 2010. **48**.

290. Hewish, M.J., Birch, D.F., and Fairley, K.F., *Ureaplasma urealyticum serotypes in urinary tract disease.* Journal of clinical microbiology, 1986. **23**(1): p. 149-154.

291. Naessens, A., et al., *Serotypes of Ureaplasma urealyticum isolated from normal pregnant women and patients with pregnancy complications.* Journal of clinical microbiology, 1988. **26**(2): p. 319-322.

292. Knox, C.L. and Timms, P., *Comparison of PCR, nested PCR, and random amplified polymorphic DNA PCR for detection and typing of Ureaplasma urealyticum in specimens from pregnant women.* Journal of clinical microbiology, 1998. **36**(10): p. 3032-3039.

293. Waites, K.B., et al., *Congenital and opportunistic infections: Ureaplasma species and Mycoplasma hominis.* Semin Fetal Neonatal Med, 2009. **14**.

294. Murphy, E.C. and Frick, I.M., *Gram-positive anaerobic cocci--commensals and opportunistic pathogens.* FEMS Microbiol Rev, 2013. **37**(4): p. 520-53.

295. Levy, P.Y., et al., *Finegoldia magna: a forgotten pathogen in prosthetic joint infection rediscovered by molecular biology.* Clin Infect Dis, 2009. **49**(8): p. 1244-7.

296. Holst, H., et al., *Detection of anaerobic prosthetic joint infection by PCR and DNA sequencing--a case report.* Acta Orthop, 2008. **79**(4): p. 568-70.

297. Ezaki, T., et al., *Proposal of the genera Anaerococcus gen. nov., Peptoniphilus gen. nov. and Gallicola gen. nov. for members of the genus Peptostreptococcus.* Int J Syst Evol Microbiol, 2001. **51**(Pt 4): p. 1521-1528.

298. Ezaki, T., et al., *Transfer of Peptococcus indolicus, Peptococcus asaccharolyticus, Peptococcus prevotii, and Peptococcus magnus to the Genus Peptostreptococcus and Proposal of Peptostreptococcus tetradius sp. nov.* International Journal of Systematic and Evolutionary Microbiology, 1983. **33**(4): p. 683-698.

299. Pagnier, I., et al., *Non-contiguous finished genome sequence and description of Anaerococcus provenciensis sp. nov.* Stand Genomic Sci, 2014. **9**(3): p. 1198-210.

300. Bowler, P.G. and Davies, B.J., *The microbiology of infected and noninfected leg ulcers.* Int J Dermatol, 1999. **38**(8): p. 573-8.

301. Shilnikova, I.I. and Dmitrieva, N.V., *Evaluation of Antibiotic Susceptibility of Gram-Positive Anaerobic Cocci Isolated from Cancer Patients of the N. N. Blokhin Russian Cancer Research Center.* Journal of pathogens, 2015. **2015**: p. 648134-648134.

302. Diop, K., et al., *'Anaerococcus mediterraneensis' sp. nov., a new species isolated from human female genital tract.* New microbes and new infections, 2017. **17**: p. 75-76.

303. Lee, S.E., et al., *The intensity of the fetal inflammatory response in intraamniotic inflammation with and without microbial invasion of the amniotic cavity.* Am J Obstet Gynecol, 2007. **197**(3): p. 294 e1-6.

304. Berg, T.G., et al., *Ureaplasma/Mycoplasma-infected amniotic fluid: pregnancy outcome in treated and nontreated patients.* J Perinatol, 1999. **19**.

305. Harwick, H.J., Iuppa, J.B., and Fekety, F.R., *Microorganisms and amniotic fluid.* Obstet Gynecol, 1969. **33**.

306. Prevedourakis, C., Papadimitriou, G., and Ioannidou, A., *Isolation of pathogenic bacteria in the amniotic fluid during pregnancy and labor.* Am J Obstet Gynecol, 1970. **106**.

307. Prevedourakis, C.N., Strigou-Charalabis, E., and Kaskarelis, D.B., *Bacterial invasion of amniotic cavity during pregnancy and labor.* Obstet Gynecol, 1971. **37**.

ABSTRACT

EVALUATING HISTORICAL PARADIGMS OF STERILITY IN PERINATAL MICROBIOLOGY AND RAMIFICATIONS FOR PREGNANCY OUTCOMES

by

JONATHAN MURRAY GREENBERG

August 2020

Advisor: Dr. Kevin R. Theis

Major: Biochemistry, Microbiology, and Immunology

Degree: Doctor of Philosophy

Next-generation sequencing technologies, especially 16S rRNA gene and metagenomic sequencing have allowed investigations of low microbial biomass tissues of the human body. While these sequencing methodologies have provided large amounts of reliable data for higher microbial biomass sites, such as the mouth, intestine, and vagina, tissues of low microbial biomass sites are subject to specific caveats that were not appropriately considered in early investigations of these sites. Low microbial biomass sites of particular interest have included those of the reproductive and urinary systems. Utilization of DNA sequencing methodologies have allowed researchers to challenge existing paradigms of sterility around these sites that were historically considered sterile, including but not limited to the placenta, the endometrium, and the bladder. While a thorough and complete understanding of the microbial signals in urogenital compartments is necessary for the best patient care and treatment, premature conclusions that redefine historical paradigms can have harmful consequences on patient health, especially for

pregnant women with whom microorganisms have been associated with multiple adverse pregnancy outcomes.

In this dissertation, I present a lack of evidence for a placental microbiota in humans using multiple modes of microbiological inquiry. Through culture, quantitative real-time PCR (qPCR), 16S rRNA gene sequencing, and metagenomics we found no evidence of bacterial signals beyond those also present in background technical controls. This work with human subjects was subsequently complemented by work in mice, in which we investigated the bacterial signals in the murine placenta and fetus, as well as multiple murine tissue control sites; we again found no consistent evidence of a placental microbiota or *in utero* colonization through multiple microbiological methodologies. Conversely, investigations of the urine of pregnant women revealed evidence of a low abundance bladder microbiota. We found bacterial signals that clearly exceeded those of technical controls, suggesting that a shift in sterility paradigm for the upper urinary tract may be warranted. Specifically, through bacterial culture, qPCR, and 16S rRNA gene sequencing we found evidence of a bladder microbiota in pregnant women that showed strong variation among individuals and consisted of *Ureaplasma urealyticum* and Gram-positive anaerobic cocci. A more thorough understanding of the bladder microbiota in pregnant women across gestation will allow healthcare professionals to address urinary and bladder symptoms in a way that alleviates or prevents pregnancy complications.

This body of work provides strategies for the thorough investigation of low microbial biomass sites and demonstrates the high degree of evidence necessary to overturn classic paradigms of sterility in perinatal medicine and host biology in general.

AUTOBIOGRAPHICAL STATEMENT

As I am finishing my tenure as a graduate student, I can feel a deep-seeded excitement for what is to come. One of the hallmarks of my personality, even since I was a child, has been my inquisitive nature. There have been more times than I can count that I have felt the frustration of others from my relentless questioning, like a child asking "why does this thing do that", with a follow-up, "but, why", and another… and another. This was a time before the internet, when Googling something was not the answer to any random question a person might have. But believe me, my questions never stopped. They reached the Googling era, and my mom has used that line on me on more than a few occasions, "why don't you Google it?". My inquisitive nature is part of what brought be back to science. Having been raised by a young and quick-to-tenure Psychology professor and a nurse, who later also became a professor of nursing, I always felt an intrinsic pressure to be an intellectual of some sort. External pressures from my parents were average at most times; they were wonderfully supportive in any endeavors I chose to pursue, as long as I was doing something.

I was a lackluster high school student and endeavored on an unconventional and exaggerated route through college that eventually led me to explore psychology. I enjoyed it, but again I enjoy most things, but it wasn't until my final semesters that I re-engaged with the biological sciences, rediscovering a passion that I experienced in middle school and only briefly in high school. I was able to channel some of this passion into tutoring other students after having graduated college. In addition to tutoring, I got a job as a medical laboratory technician at a local hospital, which allowed me to see the medical and industrial sides of my re-awakened passion. These experiences culminated into my decision to pursue graduate school, one I had always assumed I'd do, but we know what they say about assumptions. When I first came to Wayne State, I had never heard a thing about the microbiome. It was serendipitous that Dr. Theis was hired the same time that I started my first semester. And I honestly have never looked back. I am excited for the future, new directions in life and my professional career. I feel an ardent initiative to take what I have learned and completed as a graduate student and apply it to new avenues of research and most certainly many new questions.