INVESTIGATING THE HUMAN MICROBIOME IN PROSTATE HEALTH AND DISEASE: EXAMINING LOW MICROBIAL BIOMASS SAMPLES

By

Kate Bowie

B.S., BA., The University of Notre Dame, 2015

DISSERTATION

Presented to the Department of Biomedical Engineering

Oregon Health & Science University

School of Medicine

in partial fulfillment of the requirements for the degree

of Doctor of Philosophy

December 2023

Acknowledgments

First and foremost, I need to thank Dr. Lisa Karstens for co-mentoring me throughout my PhD. I approached her wanting to study the human microbiome in 2019 and she immediately took me under her wing and invited me into her group. She has instilled in me scientific rigor and played an integral part in my career development. I am eternally grateful for her patience, generosity, and her uncanny ability to know when I needed a push. I am proud of the scientist I have become, and it is largely because of her.

Secondly, I need to thank Dr. Paul Spellman. He has always been in my corner and one of my biggest supporters. He constantly challenged me to think deeper and reach higher with my research and more importantly, to be proud and confident in it. The countless pep talks I've received from Dr. Spellman had a big hand in the completion of my PhD journey.

I must also thank my co-mentor Dr. Sadik Esener and CEDAR for not only allowing me, but encouraging me to pursue my project and subsequently co-mentorship. Without the CEDAR environment that encourages curiosity, I may have never pursued human microbiome research. I am also grateful for the community I've developed within CEDAR. Specifically, Stefanie, Elise, and Jose have supported me through the ups and downs of graduate school and made coming into the laboratory a joy.

Thank you to the members of my committee, Dr. Laura Heiser and Dr. Tim Nice for their expertise and guidance. My graduate work is more impactful because of their advice and direction.

Thank you to my family, Scott, Carrie, John Hunter, and Kris. They've always been there to pick up the phone and listen to me rant about failed experiments or gush about an accepted conference abstract. John Hunter and Kris have always opened their home to me when I needed a break from Portland, and for those brief moments of rest, I'm so thankful. Lastly, I am immensely grateful for the support of my partner, Ian – my number one cheerleader and collaborator.

Table of Contents

List of Tables

List of Figures

Chapter 1. An Introduction to Prostate Disease and the Human Microbiome

1.1 Prostate Disease

Symptomatic cases of enlarged prostate glands, also known as prostate disease, were first described in The Medical and Physical Journal in 1800.[1,2] Prostate disease is a broad term encompassing three main disorders: benign prostatic hyperplasia (BPH), prostate cancer (PC) and prostatitis. BPH and PC are benign and malignant neoplasms, respectively, while prostatitis constitutes the infectious or inflammatory pathologies of the prostate. Prostate issues are common, and most men will experience some type of prostate issue during their lifetime.

In the US, the prevalence of BPH in men in their 40s is 25%; however, the prevalence jumps to 50% between the ages of 51 and 60,[3] and reaches over 80% among men in their 70s.[4] Neither how nor why BPH occurs is well understood.[4] Furthermore, it is unclear why certain men experience symptoms and others do not, irrespective of prostate size. Similar to BPH, PC is also a disease of aging and is the most commonly diagnosed solid-organ cancer in men in the United States (29% of all new cancer diagnoses in men annually).[5] PC is understood to arise as a result of complex interactions between genetics and microenvironmental factors.[6] After BPH and PC, prostatitis is the third most common urinary tract disease in men and can occur at any age.[7] Prostatitis is estimated to affect 11-13% of men; however, nearly half of all men will suffer from symptoms of prostatitis at some point in their lives.[8] The National Institutes of Health (NIH) has created four classifications for prostatitis – two of the categories have a bacterial

pathology while the remaining prostatitis types are characterized by inflammation with or without symptoms.[9]

Therapy and Outcomes for Prostate Disease

Treatments for prostate disease are constantly evolving. Currently, medical therapy is generally indicated for early stage BPH; whereas, more invasive therapies are reserved for refractory cases. Similarly, prostatitis is treated with medical therapy with a very limited role for invasive procedures. In contrast, early stage PC is treated with local therapies such as surgery and radiation; while medical therapies are indicated for more advanced disease. Despite progress in medical therapies, the current "gold standard" treatment continues to be surgical intervention for BPH (i.e. partial gland removal or ablation) while non-invasive methods are preferred for PC and prostatitis. Surgery is preferred to treat BPH due to issues with current drug therapies, such as limited efficacy,[10] potential to develop resistance to the medication leading to more severe BPH symptoms,[11] as well as side-effects such as loss of sexual function and neurological and psychological problems.[12] Although surgery can be associated with risks like bleeding, strictures, ejaculatory dysfunction, and sometimes requires multiple surgeries, it has a high success rate to treat BPH and improve the patient's quality of life.[13] The specific treatments for PC and prostatitis, on the other hand, depend on disease type and severity, which also dictate patient outcomes. For instance, the five-year-survival rate for localized PC is 99%, but this drops to 32% once the cancer escapes the prostate.[5] A variety of therapies are available for men with localized PC, including active monitoring, radiation, androgen-deprivation

therapy, and surgery, while androgen-deprivation therapy is the main-stay treatment for men with metastatic PC.[14] Infectious prostatitis is typically first treated with antibiotic therapy, however the infection could also be fungal or viral, influencing treatment.[15] In contrast, inflammatory (non-infectious) prostatitis which can cause the prostate gland to swell, is treated with NSAIDS and behavioral modification.

Lower Urinary Tract Symptoms (LUTS) in Prostate Disease

Despite prostate diseases having separate pathologies and etiologies, they share common symptoms which makes distinguishing between malignant (PC) and benign disease (BPH and prostatitis) a major clinical challenge. The primary symptoms of BPH, PC, and prostatitis are lower urinary tract symptoms (LUTS), although pain and blood in the urine can also occur.[6,16–18] Due to the location of the prostate and its envolopement of the urethra (**Figure 1.1**), enlargement or inflammation of the prostate can impact the bladder and even obstruct the urethra. LUTS describes the specific problems with storage and voiding of urine, and LUTS severity is assessed using the International Prostate Symptom Score (IPSS) self-administered questionnaire. The IPSS scores symptoms such as weak stream, nocturia, urinary frequency, and incomplete bladder emptying, to name a few.[19] In the US, treatments for LUTS have an estimated annual cost of $3.9 billion, demonstrating the large economic burden of LUTS.[20,21] Outside of prostate disease, LUTS alone exerts a negative impact on health resulting in reduced

quality of life. Since LUTS is associated with BPH, PC and prostatitis, testing to differentiate between diseases is imperative.

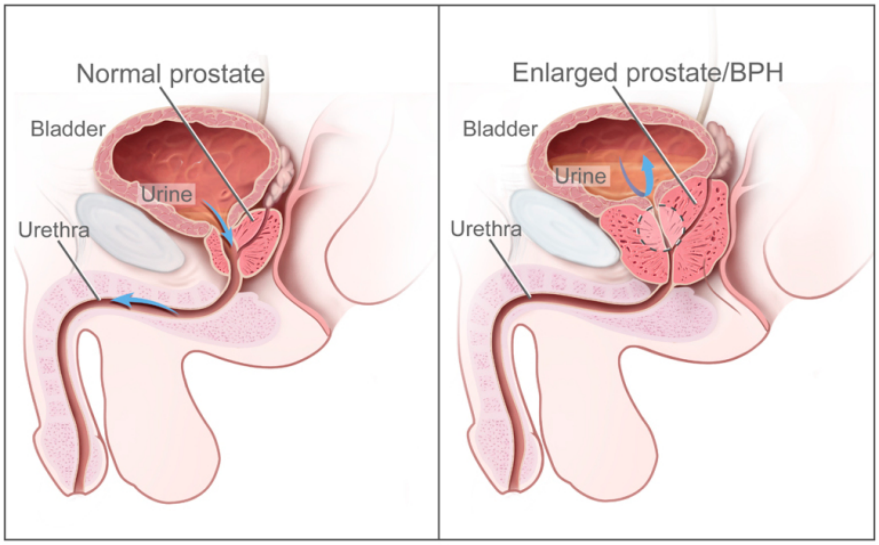


Figure 1.1. Schematic of the urogenital tract showing how the prostate interacts with the bladder.
Prostate disease, such as BPH (right, an enlarged prostate) can obstruct the urethra and press into the bladder causing lower urinary tract symptoms. Figure reused with permission.[22]

Current Clinical Practices in Diagnosing Prostate Disease

Current standard tests to differentiate BPH from PC and prostatitis include a physical exam, urinalysis, and blood test. The physical exam involves a digital rectal exam (DRE) which is done to assess prostate size and any presence of unusual nodules or masses, that could be an indicator of cancer. A voided urine sample is necessary for urinalysis when the patient presents with hematuria or if an infection is suspected, and often blood tests are recommended to measure levels of prostate specific antigen (PSA). PSA is a

serine protease produced by the epithelium of both the normal and malignant prostate. A small amount of PSA enters circulation under normal conditions and is often increased in prostate disease. The PSA test for PC was first introduced in 1990,[23] however it's utility for PC screening is still under discussion.[24,25] PSA correlates with prostate size;[26] however, PSA can be elevated due to inflammation or from infection,[27] thus it has been found to be abnormal in BPH and in 70% of prostatitis cases.[28] Importantly, PSA is not always useful for differentiating BPH from PC, as a quarter of men with BPH will have increased PSA at levels indicating PC, and will often undergo an unnecessary prostate needle biopsy.[29] The PSA test can also have false positives for clinically meanigful PC, which result in overdiagnosis and overtreatment as well as causing fear and lowering quality of life.[25]

Advancements in Understanding Prostate Disease

Monitoring prostate disease is challenging and reliance on invasive biopsies and PSA has proven to be problematic. Researchers are now looking into non-invasive methods to study and surveil prostate disease by utilizing liquid biopsy methods. Specifically, studies have investigated extracellular vesicles and circulating tumor cells isolated from the blood and urine of prostate cancer patients to use as potential biomarkers.[30–32] In addition to circulating molecules, researchers are also exploring if the quantity of cell-free circulating DNA can be useful in determining malignant from benign prostate disease.[33] There has also been potential in combining cell-free DNA assays with PSA to improve diagnostic capability.[34] Outside of human-

associated biomarkers, recent research has started to look to bacteria as a means to better understand and possibly even diagnose prostate disease.

Although prostatitis is mainly associated with inflammation often resulting in swelling, studies have shown inflammation in the prostate playing a potential role in the development of BPH and PC as well.[35,36] Researchers believe bacterial infections and even bacterial communities residing in prostate tissue are a potential catalyst for the inflammatory environment, which is also thought to spur changes associated with cancer.[37,38] Studies are still uncovering the microbial associations with prostate disease and how it relates to the human microbiome.

1.2 The Human Microbiome

The human microbiome is the collection of archaea, fungi, virus, and bacteria that inhabit the human body. In fact, the ratio of human cells to bacterial cells in the body is approximately 1:1.[39] The human body has many established microbiomes in different areas, including microbial communities that inhabit the gut, oral cavity, lungs, urogenital tract, skin, etc. (**Figure 1.2a**).[40] Although a microbiome consists of many microbial members, human studies have primarily examined bacteria and have focused on the gut microbiome. The most common phyla found in the human body are Firmicutes, Actinobacteria, Proteobacteria, and Bacteroidetes (**Figure 1.2a**), however they have different roles depending on which microbiome they are associated with. In the gut, the bacteria are highly personalized and have a commensal, beneficial, or pathogenic relationship to the host dictated by a

complex relationship between the diversity of the bacteria and their relative abundances.[41] The microbes of the gut are impacted by many factors such as genetics, age, gender, diet, medications, and environmental features like geographic location of the host (**Figure 1.3**).[42] Together, when in symbiosis with the host, the gut microbiota play major roles in human health such as shaping the intestinal epithelium, protecting against pathogens, and regulating host immunity.[41]

The immune response, and consequently disease progression, has been shown to vary drastically depending on which bacteria are present in the gut.[41,43,44] For example, *Helicobacter pylori* (*H. pylori*) is a known pathogen which causes gastric cancer, however most individuals with an *H. pylori* infection will not develop disease.[44] Additionally, co-infection of *H. pylori* with other microbes can tip the balance towards or away from cancer development, again highlighting the importance of the composition of microbiota in a community.[44] As a result, the pathogenesis of dysregulated microbes in the gut have been well-studied in relation to intestinal tract diseases and have been gaining traction as a unique factor in many other diseases.

The gut microbiome is understood to play a role in cardiovascular diseases,[45] diabetes,[46] respiratory diseases,[47] brain disorders,[48] and even cancer (**Figure 1.2b**).[49] In the context of PC, one study demonstrated the gut microbiomes of individuals with PC had altered composition with increased *Bacteroides* compared to healthy subjects.[50] In many of these studies, correlations have been made between specific bacteria found in the gut and disease outcomes. However, as the gut microbiome field has matured, researchers are now exploring more direct links, such as distinct functional

pathways influenced by bacteria and their metabolites and how that impacts disease (**Figure 1.3**).[51,52] Work to better understand the gut microbiome and its impact on human health is ongoing; however other microbiomes have been garnering interest as well.

Similar connections seen between disease and the gut microbiome have been found in other microbial niches, like the vagina,[53] oral cavity,[54] and skin.[55] The oral microbiome in particular has shown associations with organ systems distant from the oral cavity, bringing into question how far-reaching changes in specific microbiomes are.[56] There has also been reports of microbes in internal organ tissue, such as breast and prostate tissue.[57–59] Specifically, DNA from bacteria, fungi, parasites, and viruses have been found in the prostate tissue from men suffering from both BPH and PC.[58] Studies on these other microbiomes have lagged behind the gut microbiome largely because they have low microbial biomass – meaning they have a low abundance of microbes, making them challenging to study. Recent advances in methods for investigating the human microbiome are now allowing more robust study of these low microbial biomass niches.

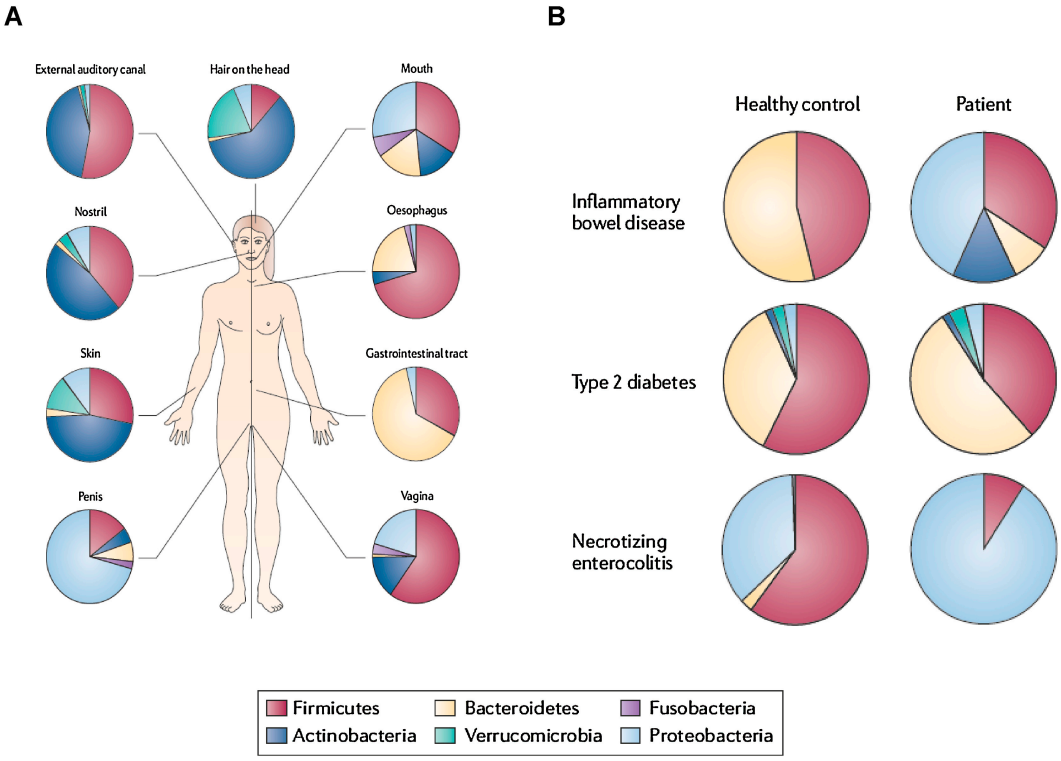


Figure 1.2. Overview of the microbiomes in the human body and how disease impacts their composition.

a) Pie charts are example representations of the different bacteria within each community in the various locations in and on the human body. **b**) Composition of the gut microbiome is altered in certain diseases like inflammatory bowel disease, type 2 diabetes, etc. Figure reused with permission.[60]

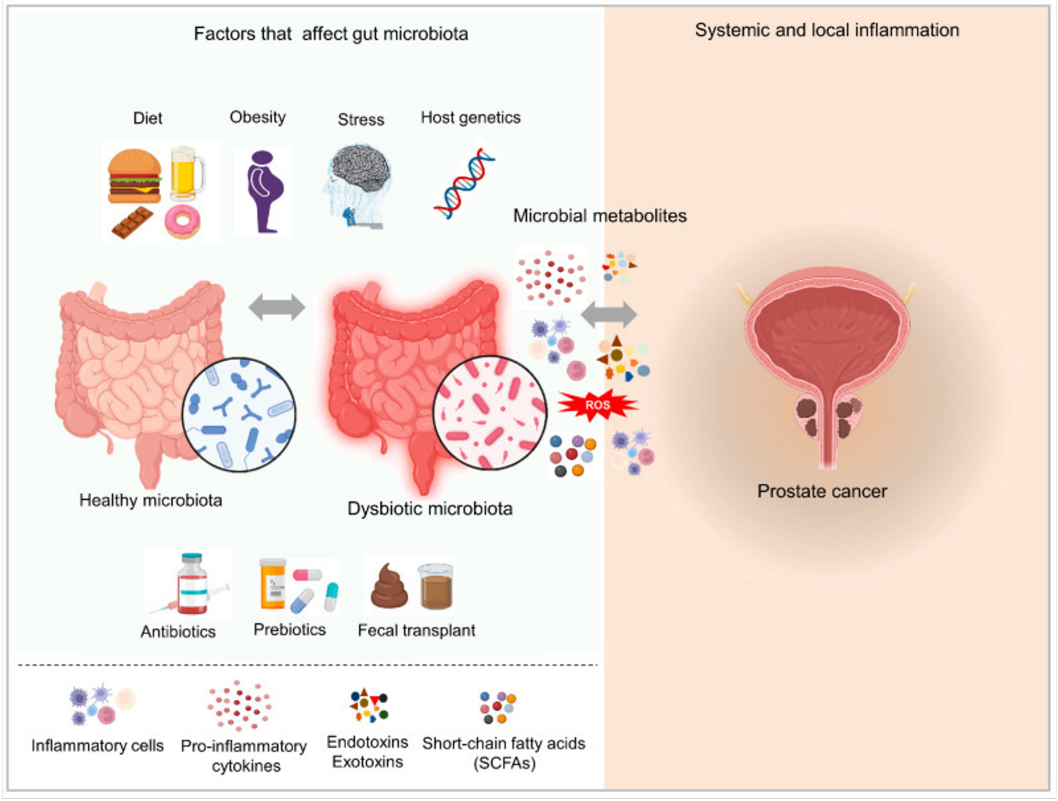


Figure 1.3. Cartoon depicting the many factors which influence the gut microbiome (diet, obesity, antibiotics, etc.) and their downstream effects.

When the gut microbiome is in a dysbiotic state, the resulting metabolites can potentially impact other organs such as the prostate. Figure reused with permission.[61]

1.3 How to Study the Microbiome

A typical workflow for studying the human microbiome is collecting a sample, extracting the DNA, sequencing the DNA, followed by data analysis (**Figure 1.4a**). Specific methods for studying the human microbiome are dependent on the sample of interest – mainly if the origin is high or low microbial biomass. Gut microbiome samples are high microbial biomass samples, with a typical concentration of 10^{11} cells per milliliter while a urinary tract microbiome sample is low microbial biomass with a typical concentration of 10^{3}-10^{5} cells per milliliter.[62] Another consideration is the biological question and what taxonomic resolution (genus vs strain-level, for example) is needed to answer that question.

DNA Sequencing

DNA sequencing is the mainstay of how researchers study the human microbiome. Specifically, Next Generation Sequencing (NGS) is the gold standard for microbiome analysis due to being high-throughput, able to examine unculturable bacteria, and increasingly affordable. Different sequencing methods include metagenome and marker gene sequencing. Metagenome sequencing captures all the DNA in a sample regardless of if it's microbial or not. This is one of the preferred methods for high microbial biomass samples, since the majority of the sequencing reads should be microbial. While this type of sequencing is more expensive than marker gene sequencing, it allows researchers to potentially reconstruct partial or even whole genomes for organisms in a sample, resulting in more precise

taxonomic resolution. Typically, species and even strain-level information about microbes can be determined using metagenome sequencing.

Targeted marker gene sequencing is often used for low microbial biomass samples and samples that have high host DNA or non-microbial DNA because it allows for the amplification of microbial DNA. In particular, sequencing the 16S rRNA (16S) gene – a 1,500 base pair (bp) gene conserved in all prokaryotes – allows researchers to resolve the different types of bacteria present in a mixed sample. The 16S gene is made up of 9 regions (V1-V9) that are each highly variable and distinct between organisms, but flanked by highly conserved regions which primers can target (**Figure 1.4b**).[63] The standard method is to use Golay primers that target the 250 bp V4 region of the 16S gene optimized for the Illumina MiSeq sequencing platform, validated by the Earth Microbiome Project and the Human Microbiome Project. This method is able to accurately describe the heterogeneity of bacterial communities,[64] typically to the family or genus level. Recently, there has been a push to sequence the full-length of the 16S gene. Full-length 16S sequencing should allow for the specificity of targeting bacteria only, while also enabling species and potentially strain-level information to be obtained from the data.

Considerations for low microbial biomass samples

Sequencing low-biomass samples is challenging as contamination can make up a larger proportion of the sequencing reads compared to a high biomass sample such as stool. This results in a lower signal-to-noise ratio for low biomass samples, thus both negative and positive controls are crucial to

establish which bacteria are from contamination.[65] To account for contamination throughout sample processing and sequencing preparation, commercial mock bacterial communities of known compositions are often diluted in a series, processed alongside the samples of interest, and sequenced to act as an in-lane control.[65] The dilution series of the mock bacterial communities are particularly useful for sequencing low biomass samples, as any genera that are not part of the expected mock community are contaminants. Other strategies include the use of spike-in controls, in which known bacteria that are distinct from what is expected in your samples of interest are added at known concentrations.[66] Spike-in controls are powerful because their input concentration is known, thus they have the added benefit of giving you abundance information about your samples. However, it requires prior knowledge of the types of bacteria expected in your samples.[66]

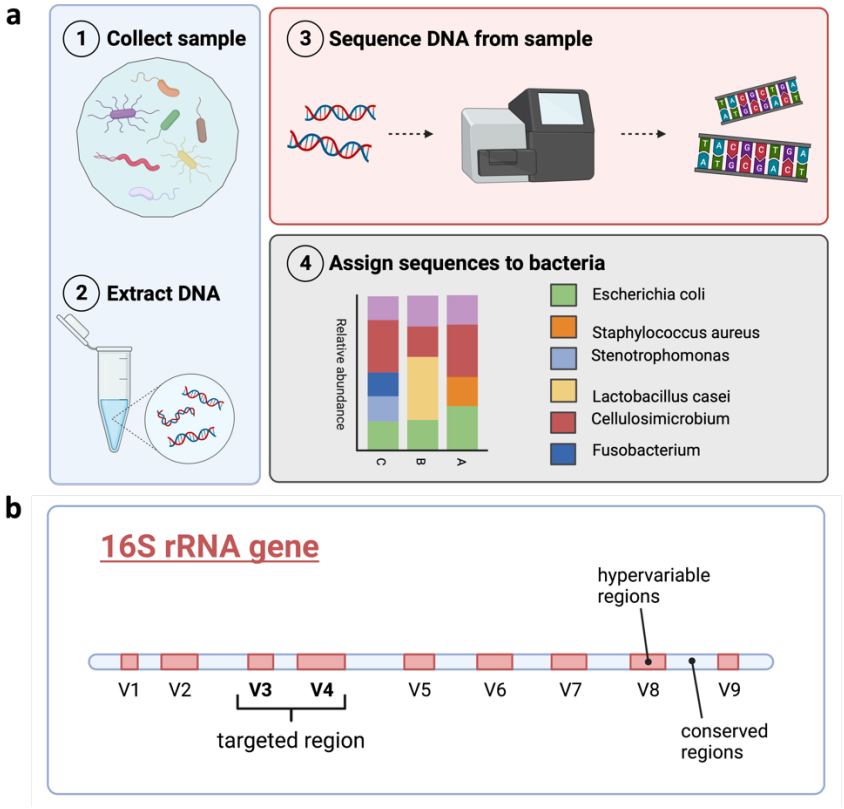


Figure 1.4. Overview of 16S rRNA gene sequencing and experimental workflow.

a) Workflow for microbiome sequencing experiments. **b)** Cartoon of the 16S rRNA gene.

Alpha and Beta Diversity Metrics:

Human microbiome analysis borrows many of its analysis metrics from ecology, such as alpha and beta diversity measures. These are measures which evaluate and describe the diversity of a microbial community, or microbiome. Alpha diversity examines within-sample diversity, or how many different microbes there are (richness) and the distribution (evenness) of those microbes in a sample.[67] There are many ways to measure alpha diversity – equating to different ways of accounting for microbial richness and evenness. Beta diversity metrics measure differences in microbial composition and abundance between samples. Certain beta diversity measures take into account the evolutionary relatedness, or phylogeny, of the microbes (UNIFRAC), while other measures rely on overlap of microbes between samples (Bray-Curtis).[68] Each alpha and beta diversity metric provides a glimpse into the diversity of the communities, and thus it is standard practice to calculate and report multiple alpha and beta diversity metrics.

1.4 New Microbial Niches

Overview of the Urinary Microbiome

As sequencing and microbiome analysis methods improve and allow for robust study of low microbial biomass samples, researchers are finding evidence of bacteria in new places. Areas of the human body previously

understood to be sterile have been found to harbor bacteria, such as in the urogenital tract and blood. Traditionally, the lower urinary tract has been considered sterile aside from the instance of an active infection. Recent research has challenged this dogma and is finding that the lower urinary tract and bladder of individuals without infection do contain microbes. This community of microbes is now known as the urinary microbiome.[69–71] The bulk of the research on the urinary microbiome has been on the female urinary microbiome, and studies have demonstrated that similar to other microbiomes in the human body in which changes are associated with diseases or disorders, changes in the microbes residing in the female urinary microbiome are distinct in women with LUTS as compared to women without LUTS.[72,73] The urinary microbiome composition in women has also shown associations with treatment efficacy for urinary urgency incontinence, a sub-category of LUTS.[74]

As disorders and diseases of the prostate can affect the bladder often causing LUTS (**Figure 1.1**), there have been recent studies of the urinary microbiome in men investigating LUTS as well as prostate disease.[75,76] Efforts to use the urinary microbiome to differentiate between BPH and PC have identified compositional differences, such as higher *Bacteroidetes, Alphaproteobacteria, Firmicutes, Lachnospiraceae, Propionicimonas* and *Sphingomonas* in men with PC compared to those with BPH.[75] However, large-scale studies (n > 150) of the male urinary microbiome have yet to be undertaken and are necessary to understand the heterogeneity of the male urinary microbiome and its potential role in prostate disease.

Understanding Microbes in Blood

Like the urogenital tract, blood was previously thought to be sterile, however microbes and microbial DNA have been detected in the blood of both healthy and diseased patients. There is disagreement over whether there are living, proliferating microbes, or only microbial DNA, in the blood of healthy people, and if the detected microbial signal is representative of a true "microbiome".[77–79] The role that microbial DNA in blood plays in disease is not well-studied, yet microbial DNA in blood has been described in type 2 diabetes, celiac disease, liver cirrhosis, chronic kidney disease, and cancer, among many others.[80–84] The microbiota in the blood are hypothesized to originate from the gut, and a preclinical study in patients with type 2 diabetes demonstrated a 28% overlap of the microbial DNA in blood and what's been found in their gut microbiome.[80] It is known that bacteria and bacterial DNA activate the immune system resulting in inflammation.[85] As inflammation contributes to the pathogenesis of liver cirrhosis, for example, it is no surprise that elevated levels of bacterial DNA have been detected in the blood of liver cirrhosis patients compared to healthy individuals.[84]

Knowledge of bacteria in blood and its association with prostate disease is still emerging. Poore et al. found PC-specific microbial DNA signatures in blood, yet their methods are under debate and the sample size was small.[83] Additionally, most studies on circulating microbial DNA have used plasma, and there is potential for bacteria to be in the non-plasma blood fractions. Overall, more work is needed to determine a relationship and better understand microbial DNA in blood with respect to prostate disease.

Investigating the Human Microbiome in Prostate Disease through Urine and Blood Samples

In summary, the overlap of symptoms between prostate disease and lack of precise testing can cause challenges for accurate diagnosis and thus proper treatment.[25,86] There have been efforts to develop alternatives to current prostate disease tests (DRE, PSA screening, etc.) such as utilizing the DNA and microbes found in urine and blood.[76,87,88] Understanding how microbes change in prostate disease could not only aid in more accurate diagnosis, but also could give hints to underlying mechanisms and possibly lead to prevention and treatment strategies.

The human microbiome is clinically important in human health and disease. Associations have emerged between the urinary microbiome and disease, as well as microbial DNA found in circulation and disease.[73,82,88–92] In men with prostate disease, both the prostate tissue itself and the urinary microbiome have been found to have altered microbial compositions, pointing to potential relationships between prostate disease and bacteria.[75,76,88,93] A limitation of the urinary microbiome studies, however, has been small sample size which is necessary to understand normal heterogeneity of the male urinary microbiome versus changes due to pathology. For microbial DNA in blood, a distinct microbial signature associated with PC has been discovered,[83] however this study and majority of others have relied on plasma as the sample of choice. Interrogation into other fractions of blood (red blood cells, buffy coat) have not been done and could yield new information.[83,92,94,95] We believe foundational studies

examining the urinary microbiome and blood in the context of prostate disease are needed.

Here, we investigate the urinary microbiome and circulating microbial DNA in blood and their relationship with prostate disease. First, we examined the urinary microbiome and associations with LUTS, BPH, PC, and prostatitis in a cohort of 500 community-dwelling men (**Chapter 2**). Next, we robustly evaluated microbial DNA in blood fractions (plasma, red blood cells, and buffy coat) from men with low- and high-grade PC, as well as their urinary microbiomes (**Chapter 3**). We also benchmarked a recent microbiome analysis method to enhance our data analysis pipeline (**Chapter 4**). Together, we believe this work will help lay the foundation for a deeper understanding of how prostate disease and the human microbiome can be studied using low microbial biomass samples.

Chapter 2. BMI and BPH Correlate with Urinary Microbiome Diversity and Lower Urinary Tract Symptoms in Men

Kate R. Bowie[1,2], Mark Garzotto[3,4], Eric Orwoll[5], Lisa Karstens[6,7]*

1. Department of Biomedical Engineering, Oregon Health & Science University, Portland, Oregon, USA

2. Cancer Early Detection Advanced Research (CEDAR), Knight Cancer Institute, Oregon Health & Science University, Portland, OR, USA

3. School of Medicine, Oregon Health & Science University, Portland, OR, USA,

4. Portland VA Medical Center, Portland, Oregon, USA

5. Division of Endocrinology, Diabetes, and Clinical Nutrition, Oregon Health & Science University, Portland, Oregon, USA

6. Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, Oregon, USA

7. Department of Obstetrics and Gynecology, Oregon Health & Science University, Portland, Oregon, USA

2.1 Abstract

Several studies have identified bacteria and other microbes in the bladder and lower urinary tract in the absence of infection. In women, the urinary microbiome has been associated with lower urinary tract symptoms (LUTS), however, similar studies have not been undertaken in large cohorts of men. Here we examine the urinary microbiome and its association with LUTS in a subset of 500 men aged 65 to 90 years from the Osteoporotic Fractures in Men (MrOS) study. We identified significant associations between benign prostatic hyperplasia (BPH), age, and body mass index (BMI) with several diversity metrics. Our analysis revealed complex relationships between BMI, BPH, LUTS, and alpha diversity which give insight into the intricate dynamics of the urinary microbiome. By beginning to uncover the interrelationships of BPH, BMI, LUTS, and the urinary microbiome, these results can inform future study design to better understand the heterogeneity of the male urinary microbiome.

2.2 Introduction

The bladder and urinary tract have long been considered sterile in the absence of infection; however, several studies have found evidence of a resident microbiome in both males and females.[69,96,97] The urinary microbiome encompasses the bacteria, archaea, fungi, and viruses that inhabit the bladder and urinary tract, and these microbes may play a role in diseases and disorders affecting the urogenital system. The majority of urinary microbiome studies have examined the female urinary microbiome and found there are various healthy microbiome compositions commonly dominated by *Lactobacillus, Gardnerella, Streptococcus, Staphylococcus, Corynebacterium,* and *Escherichia*.[73,98] Urinary microbiome studies in men have revealed several of the same bacteria found in women, such as *Escherichia, Lactobacillus,* and *Streptococcus*, but the male urinary microbiome can also be dominated by genera such as *Prevotella* and *Enterococcus*.[99–101]

Importantly, differences in the types of bacteria found in the bladders of women have been associated with recurrent urinary tract infections,[102] as well as lower urinary tract symptom (LUTS) such as urge incontinence (the sudden urge to urinate followed by leakage of urine).[72,103] The composition of the female urinary microbiome has also been found to be predictive of response to urgency urinary incontinence treatment, with a positive treatment response observed in patients with a less diverse microbiome.[102] These studies have identified specific associations between the urinary microbiome and urological diseases, and indicate more broadly that the urinary microbiome may play a role in human health and disease.

Despite the numerous studies undertaken in women, far less is known about the urinary microbiome in men.[12] A “male urinary microbiome” literature search in 2023 resulted in 22 primary research articles. Similar to the female urinary

microbiome, there are studies demonstrating that alterations in male urinary microbiome composition have been associated with diseases and disorders. Researchers have identified potential relationships between the male urinary microbiome and kidney stones,[104] benign prostatic hyperplasia (BPH),[88] as well as bladder and prostate cancers,[90,105] among many other disorders. Nonetheless, the urinary microbiome studies undertaken in men have had small cohorts (n < 100) and few have directly investigated the potential link between body mass index (BMI), age, and the male urinary microbiome. These associations could impact future study design and analysis. Larger studies are needed to both study and evaluate the broad applicability of these findings in males.

Additionally, although clear associations have been found between the female urinary microbiome and LUTS, the same associations have not been identified in a large cohort of men. Approximately one-third of men over the age of 50 experience moderate to severe LUTS.[106] The main cause of LUTS is thought to be BPH, which is hypothesized to cause symptoms primarily by pressing on the bladder and/or urethra; however, this may not be the underlying etiology for all LUTS in all men. Here, we evaluated the urinary microbiome from a large cohort of older community-dwelling men using urine samples from the NIH-funded Osteoporotic Fractures in Men (MrOS) study [107]. Our goals were to identify microbiome associations with clinical characteristics such as age, BMI, BPH, and general medical history, as well as identify associations with LUTS, including symptoms related to both urgency (irritation) and pushing or straining (obstruction) to void.

2.3 Results

Clinical Characteristics of Study Participants

We utilized urine samples and phenotypic data from 500 men at baseline as part of the MrOS study.[107,108] These 500 men were randomly selected from the overall cohort of 5994 MrOS participants. The average age of the participants in the current analysis was 73.0 ± 5.7 years, with an average BMI of 27.9 ± 3.8 kg/m^2. (**Table 2.1**). Participants completed a questionnaire detailing medical history, current health status, and overall quality of life. All participants provided a morning second-voided urine sample and completed a LUTS assessment using the International Prostate Symptom Score (IPSS). Just over half of participants (55.8%) had IPSS scores of 7 or lower and were designated as men with no/mild LUTS, while 44.2% of participants had scores higher than 7 and were designated as men with moderate to severe LUTS. All voided urine samples underwent 16S rRNA amplicon sequencing of the V4 region to determine urinary microbiome composition (see Methods). None of the participants reported a current urinary tract infection or recent use of antibiotics.

Grouped by LUTS status, the mean age in men with moderate to severe LUTS (73.7 years) was slightly but significantly higher (n=0.013) compared with those with mild to moderate LUTS (72.4 years). There was a significant difference ($p < 0.001$) in the prevalence of self-reported BPH between the two groups, with 60.5% of men with moderate to severe LUTS having a history of BPH, compared with 38.1% in the no to mild LUTS group. Significantly more men with moderate to severe LUTS also had a self-reported history of prostatitis (27.6%) compared with those with no/mild LUTS (17.7%, $p = 0.014$). The men with moderate to severe LUTS having both higher rates of BPH and prostatitis is expected, as an enlarged prostate

and inflammation are common risk factors for more severe LUTS.[109] There was, however, no correlation between BMI and LUTS severity, as both groups had a similar BMI (no to mild- 27.9 kg/m^2; moderate to severe- 28.0 kg/m^2) in contrast to one prior study which showed a higher incidence of LUTS in more obese patients.[106] There were also no significant differences in multiple baseline clinical characteristics, such as prevalence of diabetes, which was approximately 14% of the men in each group, or race (**Table 2.1**).

BPH, BMI, and Age Are Drivers of Male Urinary Microbiome Diversity

Of the total cohort, 25 samples were removed due to insufficient sequencing reads (<1000 reads), leaving a total of 475 samples for subsequent analysis. Like other body sites, the male urinary microbiome composition varied between individuals and is highly heterogenous (**Figure 2.S1a, Figure 2.S2**). Twenty-one phyla and 571 genera were identified in the male urinary microbiome, with 54 genera being core members (present in at least 10% of samples). Firmicutes was the most abundant phyla with a mean of 40.2% ± 25.7%, followed by Proteobacteria, Actinobacteria, and Bacteroidetes (27.4% ± 27.9%, 15.0% ± 15.4%, and 13.4% ± 15.7%, respectively **Figure 2.S1b**). In the majority of samples, either Firmicutes or Proteobacteria was the dominant phyla, consistent with prior male urinary microbiome studies.[89,96,110] The five most abundant genera identified were *Staphylococcus*, *Neisseria*, *Corynebacterium*, *Prevotella*, and *Streptococcus* (**Figure 2.S1c**).[1,21]

Table 2.1. Clinical characteristics of the men from the MrOS study with urinary microbiome data.

	Participants (n=475)	**No/mild LUTS (n=265)**	**Moderate/Severe LUTS (n=210)**	**p-value**
Age (y), mean (s.d.)	73.0 ± 5.7	72.4 (5.2)	73.7 (6.1)	0.013
BMI (kg/m²), mean (s.d.)	27.9 ± 3.8	27.9 (3.6)	28.0 (4.1)	0.693
Race/ethnicity, white (%)	420 (88.4%)	231 (87.2%)	189 (90.0%)	0.416
Diabetes (%) **No** **Missing**	69 (14.5%) 366 (77.1%) 40 (8.4%)	39 (14.7%) 207 (78.1%) 19 (7.2%)	30 (14.3%) 159 (75.7%) 21 (10.0%)	0.544
Prostatitis history (%)	105 (22.1%)	47 (17.7%)	58 (27.6%)	0.014
BPH history (%)	228 (48.0%)	101 (38.1%)	127 (60.5)	<0.001
Cancer history (%) **Other** **Prostate**	 89 (18.7%) 65 (13.7)	 47 (17.7%) 35 (13.2)	 42 (20.0%) 30 (14.3)	0.736

We next evaluated if the diversity of the male urinary microbiome was associated with specific clinical characteristics (**Table 2.S2, Figure 2.1a**). We identified significant associations between BPH and several measures of alpha diversity of the urinary microbiome (Shannon, Inverse Simpson, and Pielou indices, $p < 0.05$). We found that men with BPH tended to have increased alpha diversity when compared with men without BPH. We did not identify any significant associations between BPH and any beta diversity measures (weighted UNIFRAC, Unweighted UNIFRAC, Bray-Curtis), but all three beta diversity measures were significantly associated with both BMI and age (**Figure 2.1a**). We also did not identify any significant associations with urinary alpha diversity and age, BMI, diabetes, prostatitis, or race.

Based on the associations between BMI and beta diversity, we next compared the relative abundances of the five most abundant phyla and 10 most abundant genera between men with BMIs in the healthy, overweight, and obese ranges. We found no significant differences in relative abundances at the phylum level. However, we identified differences at the genus level for three genera (*Corynebacterium, Staphylococcus,* and *Streptococcus*, (**Figure 2.1b, Figure 2.S3**). Overall, obese men had the lowest abundance of *Corynebacterium, Staphylococcus,* and *Streptococcus*, whereas healthy men had the highest relative abundances. All three genera are known members of a healthy male urinary microbiome,[96,101] and significant differences in relative abundances of these taxa have been described in the gut microbiomes between obese individuals and those with lower BMIs.[111] We also examined the 5 most abundant phyla and 10 most abundant genera between individuals with and without BPH and of different age groups (determined by tertile). We found no significant differences with BPH (data not shown); however, the urinary microbiomes of men 76 to 90-years old had significantly higher relative

abundance of Actinobacteria compared to men 71 to 75-years-old (p = 0.03, **Figure 2.S4**).

We next determined how specific taxa in the male urinary microbiome were associated with age, BMI, BPH, and the remaining clinical characteristics reported in **Table 2.1**. To accomplish this, we employed an exploratory analysis using the machine-learning method, Hierarchical All-against-All (HAllA) association testing, which is high-sensitivity pattern discovery in large, paired multi-omic datasets.[112] HAllA detects possible associations with specific taxa and accompanying metadata using a statistical method for discovery. After applying HAllA at the phylum, family, and genus levels, we discovered a significant association between BMI and the genus *Dialister* (p = 0.026). *Dialister* has previously been reported in both the female and male urinary microbiome, as well as in the gut microbiome.[99,113] In the gut, higher abundances of *Dialister* were found in people with a higher BMI and is associated with difficulty in losing weight.[114–116] Similarly, in our data we found *Dialister* to be present in 23.2% of samples (**Figure 2.1c**), with the highest proportion in obese men, followed be overweight men, and lastly healthy weight men (37.3%, 20.9%, and 12.5% respectively, p<0.0001, **Figure 2.1c**). We also identified an association between the prevalence of diabetes and the phylum Bacteroidetes (p = 0.017), which has been shown to have a similar association in the gut microbiome, but has not been previously described in the male urinary microbiome.[117] We did not identify associations between specific taxa and age or BPH using HAllA. Although age and BPH are significantly associated with beta diversity and alpha diversity, respectively, these HAllA results do not point to any specific bacteria. This suggests that multiple bacteria in the microbiome community likely contribute to the differences seen above in diversity.

Clustering Reveals 8 "Urotypes"

After examining specific taxa and their associations with clinical characteristics of this MrOS cohort, we sought patterns in the overall urinary microbiome composition. In view of the associations between the urinary microbiome and BPH, BMI, and age, we determined if the associations between BMI, age, and BPH vary in the context of different microbiome compositions. We used Dirichlet multinomial modeling (DMM) to cluster samples with similar urinary microbiome compositions into "urotypes."[118] DMM clustering on members of the core urobiome revealed eight urotypes each dominated by a specific bacterium (**Figure 2.2a**). Each urotype contains between 19 and 89 samples. Urotypes 1, 2, and 7 were dominated by *Staphyloccocus*, while the rest were dominated by *Corynebacterium*, *Neisseria*, *Prevotella*, and *Anaerococcus*. Many clusters are dominated by a single bacterium while others are more diverse – for example cluster 5 is predominantly *Neisseria*, while cluster 4 is more diverse with high proportions of *Neisseria*, *Streptococcus*, and *Escherichia*.

a

Beta diversity measure	Age	BMI	Diabetes	Prostatitis	BPH	Race	Cancer
Weighted UNIFRAC	**0.02**	**0.009**	0.59	0.25	0.71	0.50	0.66
Unweighted UNIFRAC	**0.02**	**0.04**	0.91	0.98	0.31	0.31	0.66
Bray-Curtis	**0.004**	**0.002**	0.92	0.23	0.62	0.61	0.07

P-values determined by PERMANOVA model testing

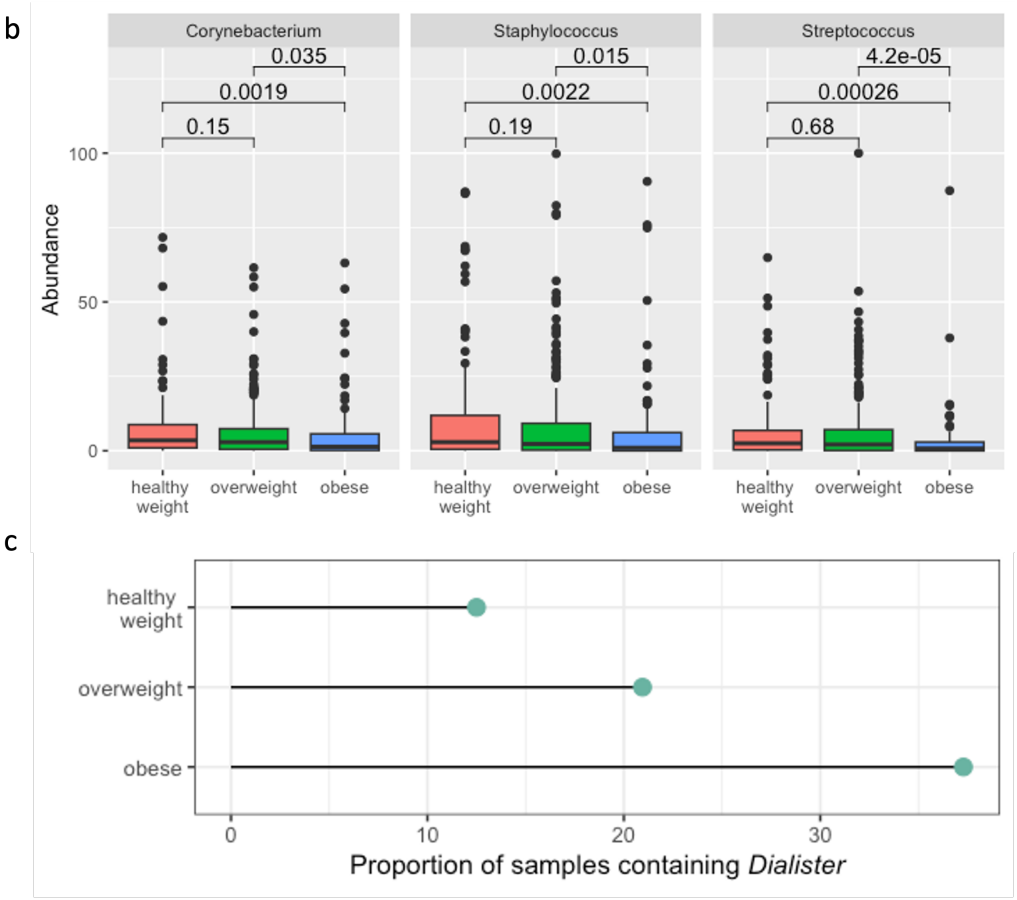


Figure 2.1: BMI is a dominant influence on beta diversity in this population. **a**) Table of p-values of PERMANOVA testing the relationship of the clinical characteristics with the beta diversity metrics of the urinary microbiome. BMI and age are both significantly associated with all three beta diversity metrics (bolded). **b**) Boxplots of relative abundance of *Corynebacterium, Staphylococcus,* and *Streptococcus* for men at healthy weight (red), overweight (green), and obese (blue) BMIs. **c**) Lollipop plot showing the percent of samples that contain *Dialister* by BMI grouping. A significantly higher proportion of obese individuals have *Dialister* in their urinary microbiomes as compared with both overweight and healthy weight individuals ($p < 0.00001$).

Next, we investigated clinical characteristics within the urotypes, and observed no significant differences in age, diabetes, prostatitis, BPH, or cancer. However, in line with our previous findings which established BMI is correlated with the composition of the male urinary microbiome, we determined the urotypes had significantly different proportions of samples belonging to healthy weight, overweight, and obese individuals ($p = 0.02$, **Figure 2.2c**). Earlier we determined a relationship between *Dialister* and the urinary microbiomes of obese men; thus, we wanted to explore that connection within these urotypes. We discovered that 88.1% of the individuals of urotype 5, which has the highest percentage of obese men (42.4%) and lowest percentage of healthy weight men (10.2%, **Figure 2.2c**), had *Dialister* in their urinary microbiomes (**Figure 2.2d**). Interestingly, urotype 8 has the second most obese men (31.6%), yet none of those men had *Dialister* in their urinary microbiomes.

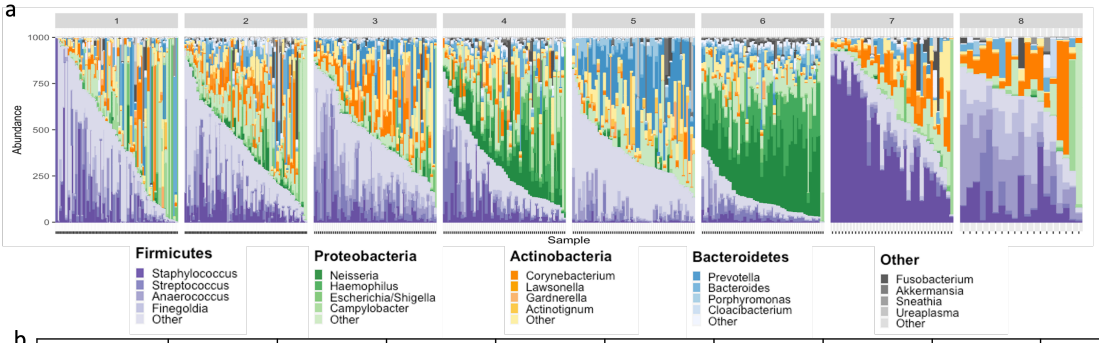
BMI, BPH and LUTS Characteristics are Associated with the Diversity of the Male Urinary Microbiome

Next, we wanted to determine if there was an association between LUTS and the male urinary microbiome. Since our results demonstrated that BPH, BMI, and age are associated with the overall male urinary microbiome composition, and previous studies have shown that BPH and BMI can also impact LUTS, we incorporated these variables into subsequent analysis.

To evaluate whether there are associations between the microbiome and LUTS, we first looked at the complexity of the samples. Men with no to mild LUTS had an average of 16,369 reads per sample (minimum 1,018; maximum 101,017),

while men with moderate to severe LUTS averaged 19,367 reads per sample (minimum 1,184; maximum 81,167; p = 0.04, **Figure 2.S5**).

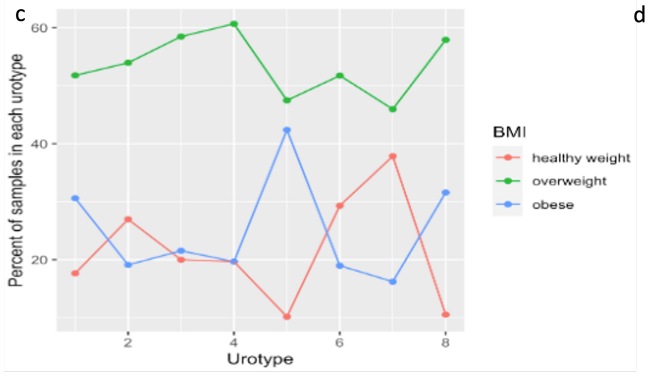
a



b

Urotype	1 n=85	2 n=89	3 n=65	4 n=61	5 n=59	6 n=58	7 n=37	8 n=19	p-value
Age, mean (s.d.)	73.3 (5.3)	72.3 (5.1)	73.2 (6.1)	73.2 (5.2)	73.8 (6.0)	71.7 (6.0)	73.1 (6.4)	73.7 (5.9)	0.48
BMI									**0.02**
Healthy	15 (17.6%)	24 (27.0%)	13 (20.0%)	12 (19.7%)	6 (10.2%)	17 (29.3%)	14 (37.8%)	2 (10.5%)	
Overweight	44 (51.8%)	48 (53.9%)	38 (58.5%)	37 (60.7%)	28 (47.5%)	30 (51.7%)	17 (45.9%)	11 (57.9%)	
Obesity	26 (30.6)	17 (19.1%)	14 (21.5%)	12 (19.7%)	25 (42.4%)	11 (19.0%)	6 (16.2%)	6 (31.6%)	
Diabetes	12 (14.1%)	9 (10.1%)	6 (9.2%)	11 (18.0%)	11 (18.6%)	10 (17.2%)	7 (18.9%)	3 (15.8%)	0.75
Prostatitis history	25 (29.4%)	20 (22.5%)	13 (20.0%)	12 (19.7%)	16 (27.1%)	10 (17.2%)	6 (16.2%)	3 (15.8%)	0.56
BPH history	44 (51.8%)	48 (53.9%)	29 (44.6%)	27 (44.3%)	28 (47.5%)	28 (48.3%)	16 (43.2%)	7 (36.8%)	0.82
Cancer									0.60
Other	10 (11.8%)	23 (25.8%)	10 (15.4%)	9 (14.8%)	12 (20.3%)	14 (24.1%)	9 (24.5%)	1 (5.3%)	
Prostate	15 (17.6%)	12 (13.5%)	13 (20.0%)	5 (8.2%)	6 (10.2%)	6 (10.3%)	4 (10.8%)	4 (21.1%)	

c



d

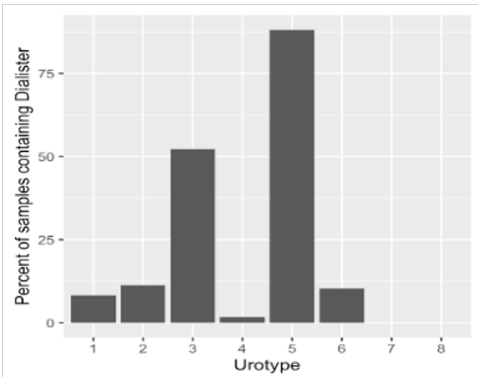
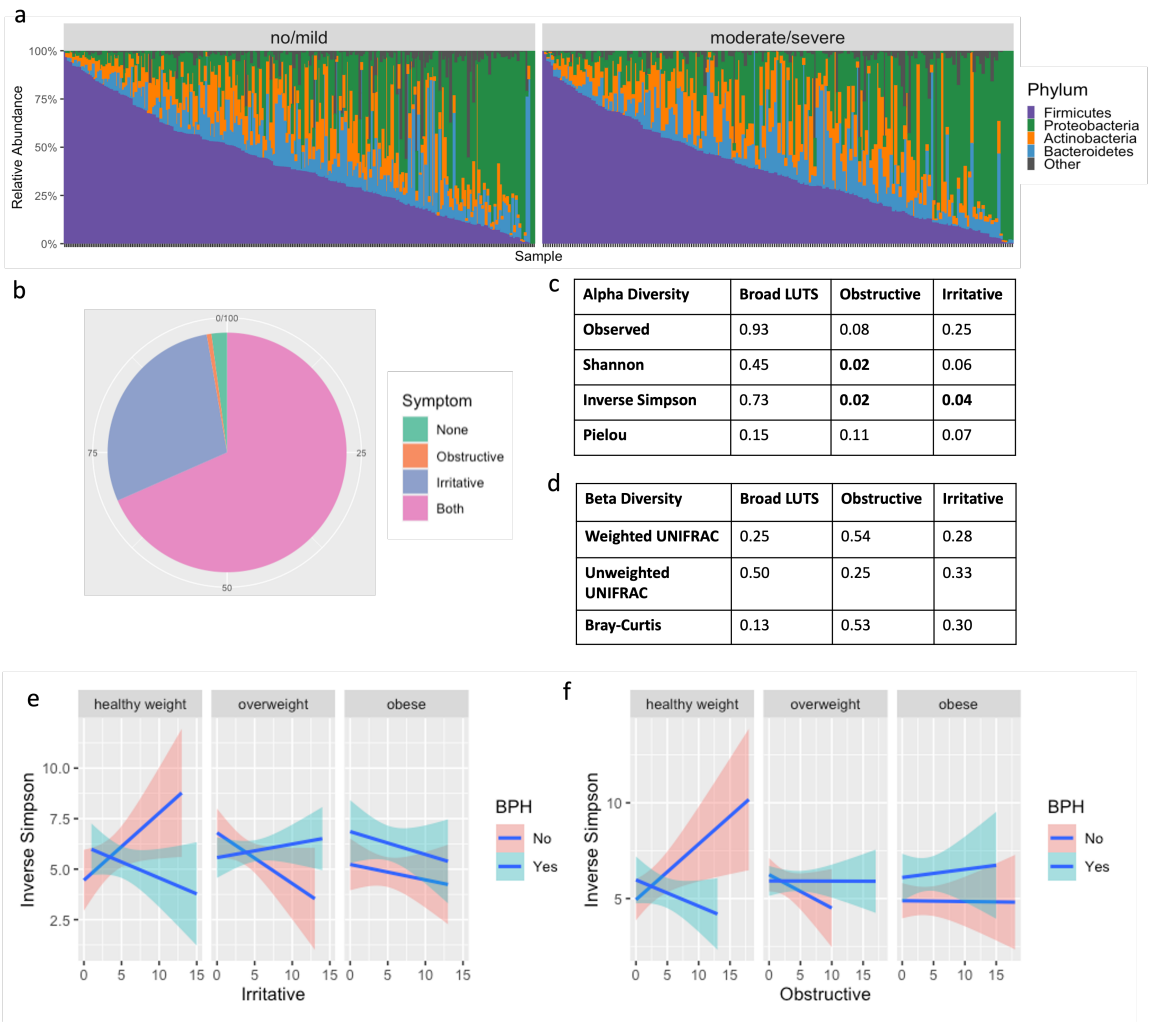


Figure 2.2. Distinct patterns of the male microbiome are present and are associated with BMI.

a) Stacked bar plots of the eight urotypes created using Dirichlet Multinomial modeling. Each urotype is dominated by specific bacteria **b**) Table summarizing the clinical characteristics for each urotype. BMI is significantly different between urotypes (p = 0.02, bolded) **c**) Traces of the percent of samples in each urotype colored by BMI group. **d**) Bar graph of the percent of Dialister in each urotype. Urotype 5 has the highest percentage of Dialister and obese men.

We then inspected the urinary microbiome compositions in men with no to mild LUTS compared with those with moderate to severe LUTS and did not find significant differences (**Figure 2.3a**). We also examined alpha diversity metrics (Kruskal-Wallis test) and beta diversity metrics (PERMANOVA analysis) of the urinary microbiome composition in these two groups, and found no significant differences were found even when adjusted for BPH, BMI, and age (**Figure 2.3c and Figure 2.3d**).



c

Alpha Diversity	Broad LUTS	Obstructive	Irritative
Observed	0.93	0.08	0.25
Shannon	0.45	**0.02**	0.06
Inverse Simpson	0.73	**0.02**	**0.04**
Pielou	0.15	0.11	0.07

d

Beta Diversity	Broad LUTS	Obstructive	Irritative
Weighted UNIFRAC	0.25	0.54	0.28
Unweighted UNIFRAC	0.50	0.25	0.33
Bray-Curtis	0.13	0.53	0.30

Figure 2.3. The alpha diversity of the male urinary microbiome is associated with irritative and obstructive LUTS when adjusted for BMI and BPH.

a) Stacked bar plot of the male urinary microbiome subset by LUTS status. **b)** Pie chart showing percentage of men in the MrOS cohort who experience no symptoms (green), obstructive symptoms only (orange), irritative symptoms only (blue), and both obstructive and irritative symptoms (pink). **c)** P-values for alpha diversity metrics and **d)** beta diversity metrics for broad LUTS (no to mild LUTS versus moderate to severe LUTS) and obstructive and irritative symptoms. Adjusted model includes interaction terms for BMI and BPH **e)** Significant interactions between BMI, BPH, and irritative or **f)** obstructive symptoms influence inverse Simpson index.

LUTS is defined broadly, encompassing both irritative and obstructive symptoms, and we hypothesized that there may be differences in the microbiome that were related to these characteristics rather than overall severity. For example, studies on LUTS and the urinary microbiome in women have found significant differences in women with stress incontinence versus urgency incontinence;[119] however, differences are not typically found when examining overall incontinence.[120] After separating the IPSS questions into those addressing irritative versus obstructive symptoms, we found the majority of men experienced some degree of both symptoms. Fewer had irritative symptoms alone, while less than 1% of men experienced only obstructive symptoms (**Figure 2.3b**). Irritative symptom scores ranged from 0 to 15, while obstructive symptoms ranged from 0 to 18.

Taking a closer look at how obstructive and irritative symptoms are associated with alpha and beta diversity, we used a Kruskal-Wallis Test and adjusted for BPH, BMI, and age. We found the Inverse Simpson alpha diversity measures significantly associated with obstructive symptoms and irritative symptoms independently (**Figure 2.3f**). Obstructive symptoms were also associated with the Shannon index, while the association with irritative symptoms was weaker, and not statistically significant (Shannon, $p = 0.06$). In PERMANOVA comparisons, we found no associations with beta diversity and obstructive nor irritative symptoms (**Figure 2.3d**). In addition to the alpha diversity, we also discovered significant interactions between irritative symptoms and BPH and BMI (**Figure 2.3e**), as well as similar results with obstructive symptoms (**Figure 2.3f**). In other words, men with BPH have varying associations between alpha diversity and symptom severity depending on BMI, and there are different associations if the individual does not have BPH. For example, in the healthy weight population, men with a higher Inverse Simpson index tended to have increased irritative symptoms if they do not

have BPH (**Figure 2.3e**). However, the converse is true if they have BPH. We see the opposite trends for men in the overweight category and found that BPH does not have the same impact in the obese group.

Enterococcaceae and Caulobacteraceae are Associated with Irritative Symptoms

In light of irritative and obstructive symptoms and their potential role in the diversity of the male urinary microbiome, we re-examined the eight urotypes by HAllA. There were no significant differences in the obstructive and irritative symptom scores between urotypes ($p>0.05$), nor when investigating broad LUTS severity.

To discover relationships with specific taxa and LUTS severity or irritative and obstructive symptoms in the older male urinary microbiome, we again employed (HAllA).[112] Supporting our previous results, no significant associations at any taxonomic resolution were identified with overall LUTS score. We investigated irritative versus obstructive symptom scores independently, which resulted in associations at the family level, between irritative symptoms and the presence of Enterococcaceae ($p = 0.002$) and Caulobacteraceae ($p < 0.001$) independently. Enterococcaceae has previously been described in the female urinary microbiome of women with mild and moderate to severe LUTS.[121] Caulobacteraceae has been found in the healthy canine urinary microbiome,[122] as well as the urinary microbiome of spontaneously tolerant kidney-transplant recipients;[123] however, no research has shown any relationship to urgency or LUTS.

We further assessed each of the questions from the American Urological Association symptom index (AUA-SI) questionnaire for a more granular analysis

and found an association at the genus level between the urge to urinate frequently and *Mobiluncus* ($p < 0.01$). *Mobiluncus* is primarily found in post-menopausal women and women with bacterial vaginosis and has not previously been reported to persist in the male urinary microbiome.[124,125]

2.4 Discussion

Our analysis of a relatively large cohort of older community-dwelling men revealed that the male urinary microbiome is heterogeneous and exhibits a great deal of inter-individuality in the present microbes. We found that the male urinary microbiome is dominated by *Staphylococcus, Corynebacterium, Prevotella,* and *Finegoldia*, which supports previous research (**Figure 2.S1c**).[96,126] We also identified complex relationships between the male urinary microbiome and BMI, BPH, age, and LUTS. While previous studies have found age-specific compositional differences in the urinary microbiome regardless of gender,[69] associations between BMI and the urinary microbiome have been established only in women.[119]

A major finding of our study is the significant association between BMI and the male urinary microbiome composition (**Figure 2.1a**). Similar associations have been identified in other microbiome communities in the human body[127] and in the urinary microbiome in women, [119] but have not yet been reported in the male urinary microbiome. This means that men with similar BMIs have similar urinary microbiome compositions, compared to men with different BMIs regardless of age, BPH, and medical history.

We found that BMI is linked to specific genera as well as to the overall microbiome composition. Using a Kruskal Wallis rank sum test, the relative abundances of *Corynebacterium*, *Staphylococcus*, and *Streptococcus* were significantly

different in healthy weight, overweight, and obese men; obese men had a decreased abundance of all three genera, whereas healthy weight men had an increased abundance (**Figure 2.1b**). Since these genera are typically found in a healthy urinary microbiome, and BMI is significantly associated with beta diversity, we can hypothesize that the lower relative abundance of these three genera allows other bacteria, like *Dialister*, to create a niche in those communities. A machine-learning approach, HAllA, revealed a significant relationship between the abundance of *Dialister* and higher BMI (**Figure 2.1c**), an association which has not been previously reported in the context of the urinary microbiome.[114–116] In the gut, many studies have shown *Dialister* in individuals with higher BMIs and who have difficulty losing weight.[111,115,127,128] One study surmised that *Dialister* aggravates the host inflammatory response and insulin resistance by releasing more lipopolysaccharides,[115] which are known to be an important feature in metabolic disease and weight gain.[129] Interestingly, in the DMM clustering that revealed eight urotypes (**Figure 2.2**), urotype 5 contained the highest percentage of obese men and the highest percentage of men with *Dialister*. Although this mechanism has not been studied in the urinary microbiome, a similar one could be at play. Interestingly, in the DMM clustering that revealed eight urotypes (**Figure 2.2**), in the second highest cluster of obese men (urotype 7) *Dialister* was not detected. We believe this is something to investigate further and examine if other bacterial communities in the urinary microbiome could have an association with BMI.

We also observed that men with BPH tended to have a higher alpha diversity in their urinary microbiome as compared to men without (**Table 2.S2**) – a finding that had previously not been reported in studies.[130] BPH has previously been associated with significant changes in the urinary microbiota and in the microbes of the prostate tissue itself.[75,130] In rats, BPH has been associated with the beta diversity

of the gut microbiome; however, no such studies have been undertaken in humans.[131]

We did not identify significant associations between overall LUTS score and the urinary microbiome composition in men. Previous work has shown an association between more severe LUTS and the following genera: *Haemophilus*, *Staphylococcus, Dolosigranulum, Listeria, Phascolarctobacterium, Enhydrobacter, Ruminococcus, Bacillus, Faecalibacterium, and Finegoldia*.[130] We did not find *Listeria* nor *Ruminococcus* in our dataset, and there were no associations between the other genera and LUTS severity. However, we believe the overall IPSS severity score may not allow for the detection of associations with specific causes of urinary symptoms (e.g., irritative and obstructive symptoms), which could have different etiologies. We considered the hypothesis that individual elements of the LUTS symptoms score could have different associations with the microbiome. For instance, irritative symptoms could have more inflammatory bacteria, while obstructive symptoms could be the result of bacteria more likely to create a biofilm which could create some blockage while voiding. Or perhaps these bacteria could lead to increase tissue growth and swelling which could cause urinary obstruction. In support of this idea, we did find significant associations between alpha diversity metrics and specific obstructive and irritative LUTS symptoms when the analyses were adjusted for BMI, BPH, and age. The inverse Simpson index was significantly associated with both obstructive and irritative symptoms; however, different trends were observed depending on BMI grouping and BPH (**Figure 2.3e-f**). Similar to what was discussed earlier with *Dialister* and BMI, these results suggest that there could be different mechanisms causing irritative or obstructive symptoms in obese men as compared to overweight men with or without BPH.

HAllA unveiled associations at the family level between Enterococcaceae and Caulobacteracaea with irritative symptoms; however, no associations were found at the phylum or genus level, nor any associations with obstructive symptoms. The HAllA analysis was restricted to 54 of the 571 genera, so this may have contributed to our limited findings. More work is needed to determine which taxa are important for irritative and obstructive symptoms; however, the heterogeneity of bacteria between individuals and other characteristics of the microbiome may not yield consensus. In other words, biological functions may be more homogenous and may be more indicative of phenotype or symptoms, despite being caused by a number of different bacteria.

The large sample size of this study allowed us to investigate broad trends across the urinary microbiome of older community-dwelling men. Although we had a rich dataset, there were some limitations that may have affected our analysis. Elements of the phenotypic data (e.g., history of diabetes and cancer) were self-reported and thus subject to possible errors. The samples collected were voided urine, which have been reported to not represent only the male bladder microbiome, but instead be a mixture of the bladder and urethral communities.[89] Also, as with any 16S rRNA sequencing study, analysis was limited to the genus level, and resolution to the species level may improve findings. For example, there are many different species of *Dialister*. Finally, the LUTS score is based on a set of symptoms that may be the result of multiple biological pathways, and thus difficult to disentangle.

In summary, the male urinary microbiome is complex and challenging to investigate. There is considerable individual variation in the urinary microbiome in men, and that variation may be related to the presence and character of LUTS. Our study indicates that considering the specific type of LUTS is important in

understanding the contribution of urinary microbiota to urinary disorders in men. We also revealed there is an intricate connection between BMI, BPH, urinary microbiome diversity, and severity of irritative and obstructive symptoms. More studies are needed to further investigate these covariates, and we believe future urinary microbiome studies in men should heavily consider BMI and BPH in study designs and analyses.

2.5 Methods

Study Population

We acquired samples and data from 500 randomly selected participants from the prospective NIH-funded MrOS study. The MrOS study recruited a total of 5,994 men between the ages of 65 and 100 years from 6 clinical sites in the United States to assess risk factors for fracture and other conditions related to aging. The cohort and recruitment methods have been previously described.[132] At baseline, participants were at least 65 years of age, able to consent, walked without assistance of another person, and did not have bi-lateral hip replacement or any condition that in the judgment of the site investigator would likely impair participation in the study.[133] The Institutional Review Boards at all sites reviewed and approved the study, and all participants provided written informed consent. Enrolled participants completed a series of medical questionnaires, including medical history of prostatitis and BPH and the AUA-SI questionnaire, and provided specimens such as blood and urine which were banked for future research. Participants were recruited between 2000 and 2002 and followed longitudinally. Morning, second-voided

urine specimens were collected in sterile containers and frozen at -80 °C for future analyses.

DNA Isolation and Sequencing

Bacterial DNA was isolated from 4-mL aliquots of urine specimens. The V4 region of the 16S rRNA gene was selectively amplified to evaluate microbial composition of the urine. DNA from the 500 urine samples was submitted to Baylor University for paired-end 250 base-pair sequencing using Illumina MiSeq by Dr. Nadim Ajami.

Classification of LUTS

LUTS was assessed using the IPSS. The IPSS is a seven-question examination that evaluates both irritative and obstructive aspects of LUTS. Each question is given a score from zero to five in terms of how severe that specific symptom is, and the scores for all questions are totaled. No/mild LUTS were assigned to scores $0 \leq 7$, while moderate/severe LUTS corresponds to scores greater than 7. LUTS was then divided into irritative versus obstructive symptoms for sub-analyses based upon the IPSS questionnaire. Scores from questions 1, 3, 4, and 6 were totaled for an overall obstructive symptom score, and the scores from the remaining questions made up the irritative score.

Bioinformatics and Statistical Analyses

Raw sequences were processed into amplicon sequence variants (ASVs) using DADA2. The RDP Classifier was used to map the ASVs to the SILVA 128 16S rRNA

reference set for taxonomic identification. The 500 samples were rarefied to 1000 reads, which removed 25 samples and left 475 samples for downstream analyses. All subsequent analyses were done on ASVs agglomerated at the genus level in R, except for application of HAllA which uses Python.[112] Data were further processed using phyloseq (version 1.42.0) and visualized using *microshades* (version 1.10).[134] The *Vegan* R package version 2.6.4 and *rstatix* version 0.7.2 were used for all statistical analyses.

Testing for Associations Between Taxa and Clinical Characteristics

We implemented HAllA version 0.8.20 as an exploratory analysis to investigate associations between individual taxa and the clinical characteristics. HAllA was carried out separately at the genus, family, and phylum levels, and we exported both our metadata and ASV tables to text files. The metadata included the clinical variables of Table 1, and separate ASV tables were generated for each taxonomic level.

We also looked at relative abundance of the top 5 phyla and top 10 genera for BMI, BPH, and age. A Kruskal-Wallis test was used to test for a significant association between relative abundance and clinical characteristic of interest, and then a pairwise Wilcoxon Rank Sum test with false discovery rate correction.

Clustering of Samples by Microbiome Composition

Treating each urinary microbiome as a community, we decided to cluster based on these communities and then look for associations with clinical characteristics and LUTS. We used the *DirichletMultinomial* version 1.40.0 for DMM clustering on the

"core" taxa, or the bacteria present in at least 20% of samples, to improve processing speed. We evaluated the model fit for two through 12 clusters using the Laplace approximation and chose eight clusters based on a global minimum (**Figure 2.S6**)[118] and found that qualitatively eight clusters exemplified eight different community types dominated by different taxa (**Figure 2.2a**). The samples were separated into eight clusters based on the core taxa, and the bacterial communities in each cluster were tested for associations with clinical characteristics using the tableone R package (version 0.13.2).

Data Availability

Data will be released through MrOS online (https://mrosonline.ucsf.edu/) upon publication.

Author Contributions

E.O. designed the study and assisted with data access and acquisition. M.G. guided data analysis and interpretation with clinical input. L.K. designed and guided the data analysis, and wrote the manuscript. K.B. analyzed the data and wrote the manuscript. All authors contributed to interpretation of results and editing the manuscript.

2.6 Acknowledgements

The Osteoporotic Fractures in Men (MrOS) Study is supported by National Institutes of Health funding. The National Institute on Aging (NIA) and the National Center for Advancing Translational Sciences (NCATS) provide support under the following award numbers: R01 AG066671 and UL1 TR000128. This project was additionally supported by NIH/NIDDK K01DK116706 (LK) and the Cancer Early Detection Advanced Research Center at Oregon Health & Science University's Knight Cancer Institute (KB).

2.7 Supplemental Figures and Tables

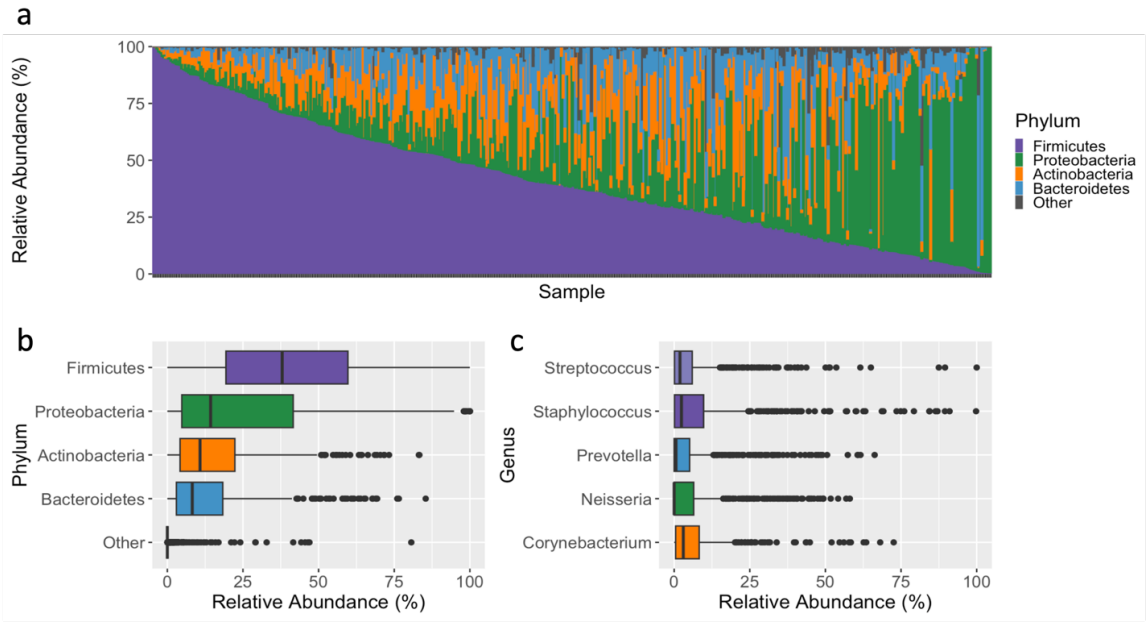


Figure 2.S1. The male urinary microbiome is highly heterogenous.
a) Stacked bar plot of the male urinary microbiome at the Phylum level ordered by Firmicutes abundance. **b**) Boxplots of the most abundant Phyla and **c**) Genera in the male urinary microbiome.

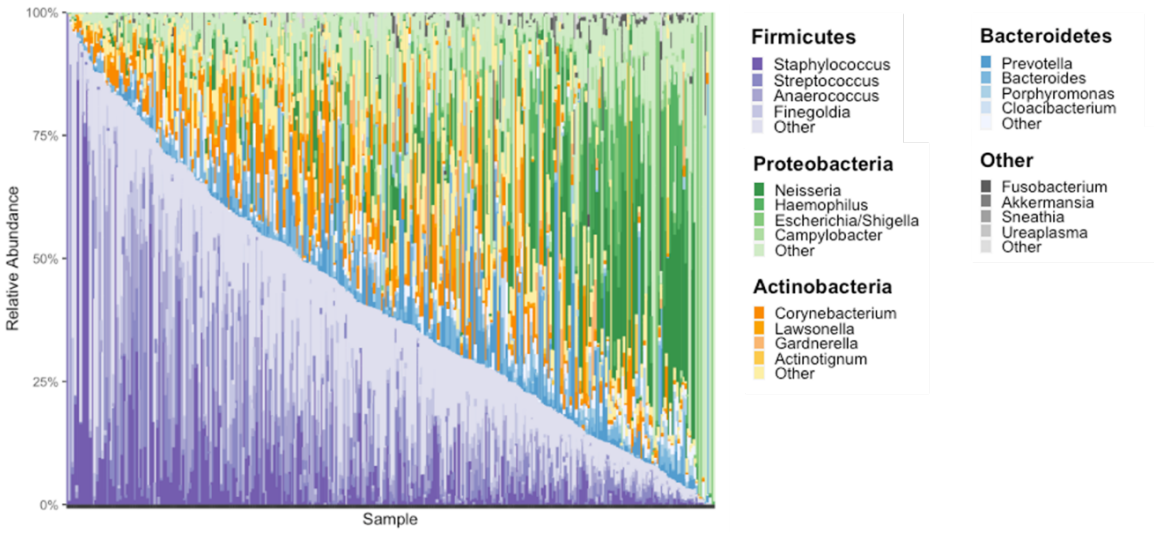


Figure 2.S2. The male urinary microbiome is highly variable demonstrated by abundance of different genera.

A stacked bar plot of the male urinary microbiome organized by Firmicutes abundance. Each phylum is assigned a color, and within each phylum the darker shades are assigned to the most abundant genera. The lightest shade of each color is the rest of the genera corresponding to that phylum. The overabundance of light shades (i.e. in Firmicutes and Proteobacteria) demonstrates that the male urinary microbiome is highly heterogenous.

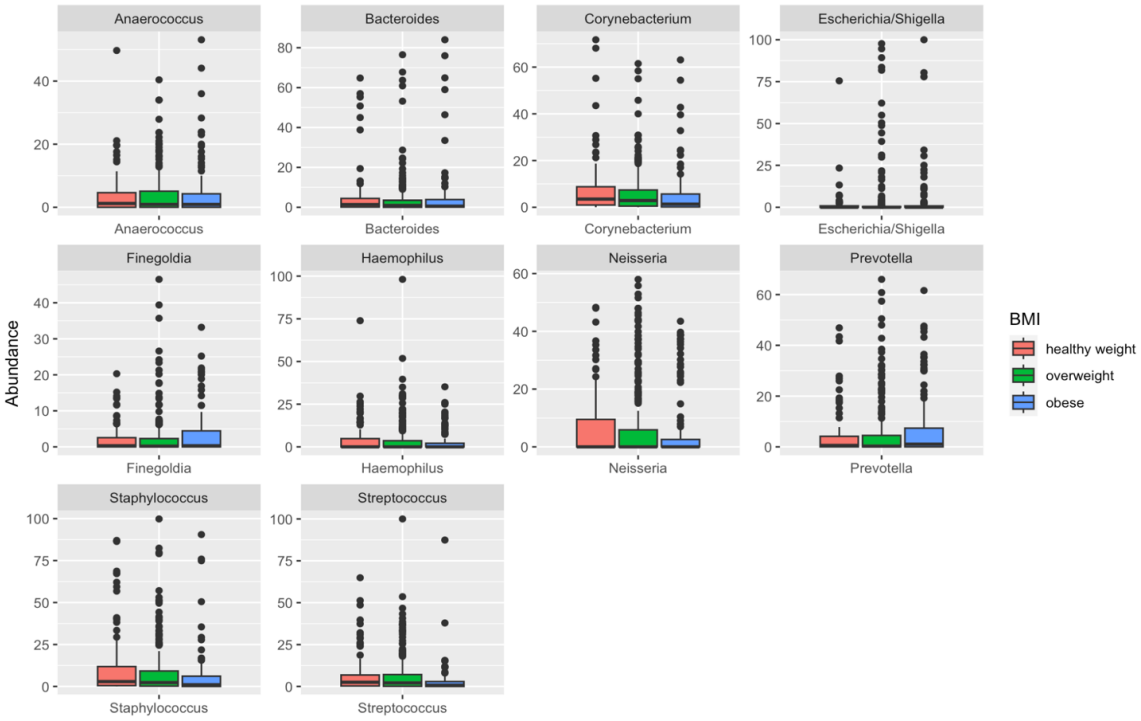


Figure 2.S3. Boxplots of the most abundant genera in the MrOS cohort by BMI group (healthy weight, overweight, and obese).

There were significant differences in abundance between groups detected for *Corynebacterium*, *Staphylococcus*, and *Streptococcus* highlighted in **Figure 2b**.

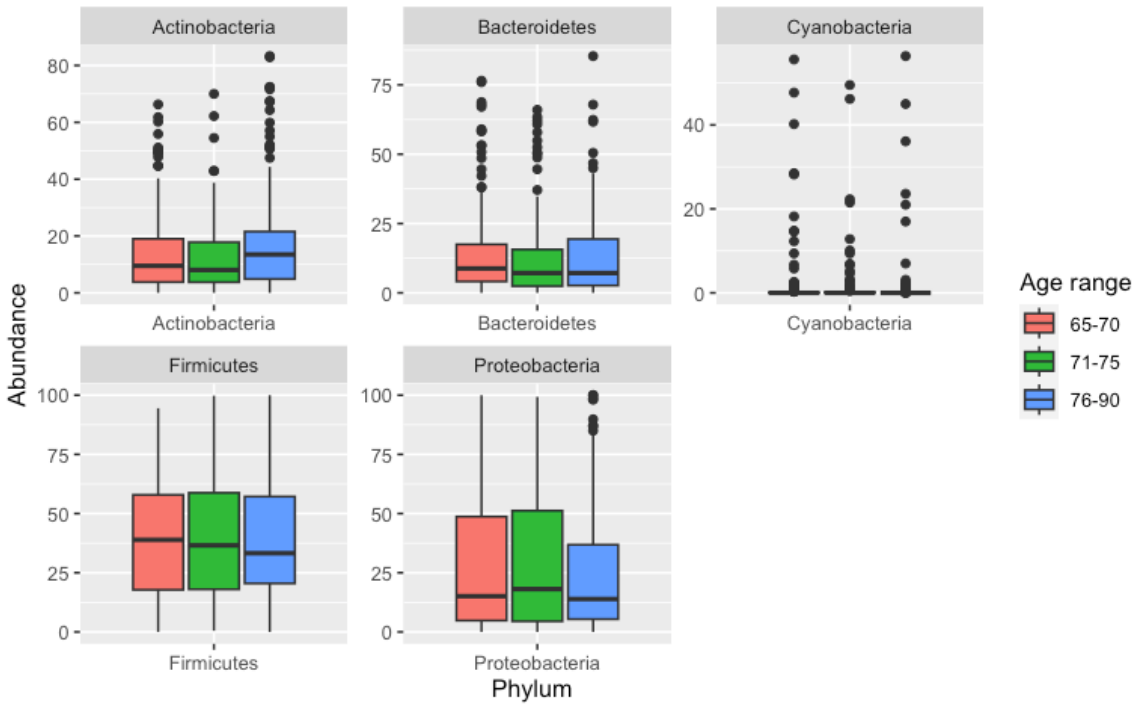


Figure 2.S4. Boxplots of relative abundances of the top five most abundance phylum by age range.

Age ranges were determined by tertiles. The relative abundance of Actinobacteria in men aged 76 to 90 is significantly higher than in men aged 71 to 75 ($p = 0.03$).

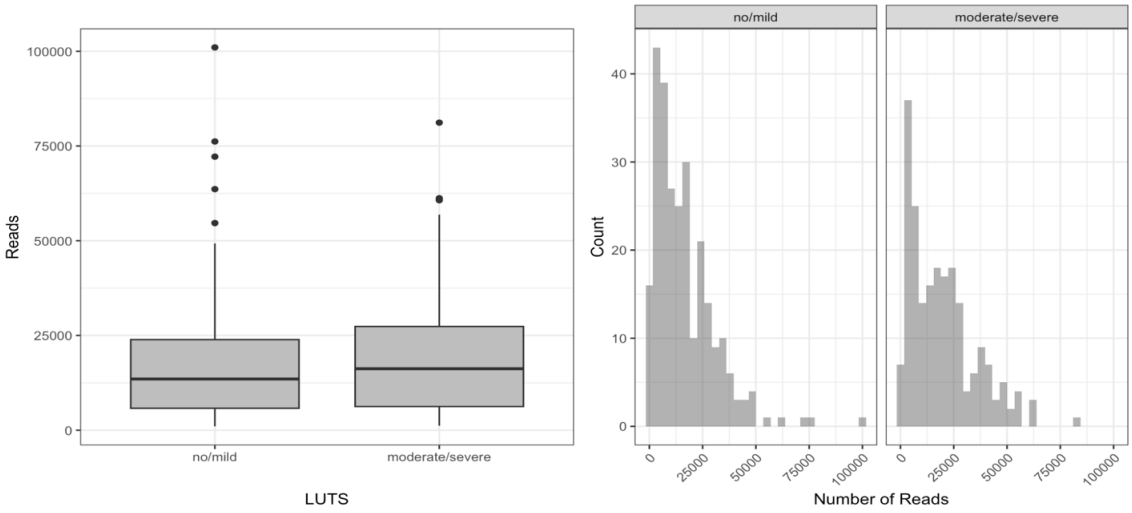


Figure 2.S5. Boxplot and histogram of the number of reads by LUTS severity. No/mild LUTS (red) has significantly less reads than moderate/severe LUTS (blue, p = 0.04).

Table 2.S1. Associations of clinical characteristics and urobiome alpha diversity.

Categorical data are reported as median [interquartile range (IQR)] of the diversity measure for the specified group, continuous variables (age) are reported as correlation coefficients. Significant associations were identified between BPH and Shannon, Inverse Simpson, and Pielou indices are significant ($p < 0.05$, bolded) determined by Wilcoxon rank sum test.

	Age		Race			BMI			
		p-value	White	Other	p-value	Healthy Weight	Overweight	Obese	p-value
Observed	-0.06	0.17	23.0 [16.0, 30.0]	26.0 [17.0, 30.0]	0.42	22.0 [16.0, 33.0]	22.0 [16.0, 29.0]	24.5 [18.0, 30.0]	0.63
Shannon	-0.06	0.22	2.1 [1.7, 2.4]	2.2 [1.7, 2.4]	0.70	2.1 [1.7, 2.4]	2.0 [1.7, 2.4]	2.0 [1.6, 2.4]	0.48
Inverse Simpson	-0.05	0.24	5.1 [3.3, 7.2]	5.0 [3.2, 7.8]	0.97	5.0 [3.0, 7.1]	5.5 [3.5, 7.4]	4.9 [3.3, 7.0]	0.31
Pielou	-0.06	0.19	0.67 [0.57, 0.75]	0.68 [0.55, 0.76]	0.92	0.66 [0.55, 0.74]	0.69 [0.58, 0.76]	0.64 [0.56, 0.74]	0.12

	Diabetes			Prostatitis		
	Yes	No	p-value	Yes	No	p-value
Observed	24.0 [18.0, 30.0]	22.0 [16.0, 30.0]	0.30	22.0 [16.0, 31.0]	23.0 [17.0, 33.0]	0.67
Shannon	2.2 [1.8, 2.4]	2.1 [1.7, 2.4]	0.40	2.1 [1.6, 2.4]	2.1 [1.7, 2.4]	0.69
Inverse Simpson	5.2 [3.5, 7.3]	5.1 [3.3, 7.1]	0.44	5.5 [3.4, 7.5]	5.1 [3.3, 7.2]	0.46
Pielou	0.66 [0.58, 0.75]	0.66 [0.56, 0.75]	0.82	0.69 [0.58, 0.75]	0.66 [0.56, 0.75]	0.39

	BPH*			Cancer			
	Yes	No	p-value	None	Other	Prostate	p-value
Observed	22.5 [16.0, 30.0]	23.0 [17.0, 29.5]	0.85	23.0 [17.0, 30.0]	23.0 [17.0, 28.0]	23.0 [15.0, 30.0]	0.83
Shannon	**2.2 [1.8, 2.4]**	**2.0 [1.7, 2.3]**	**0.04**	2.1 [1.7, 2.4]	2.1 [1.7, 2.4]	2.2 [1.7, 2.5]	0.78
Inverse Simpson	**5.8 [3.6, 7.4]**	**4.8 [3.1, 6.9]**	**0.04**	5.1 [3.2, 7.2]	5.2 [3.5, 7.1]	5.8 [3.5, 7.7]	0.55
Pielou	**0.69 [0.61, 0.75]**	**0.64 [0.55, 0.75]**	**0.03**	0.66 [0.56, 0.74]	0.67 [0.58, 0.77]	0.69 [0.58, 0.77]	0.33

Table 2.S2. Overall LUTS and obstructive and irritative symptom scores summarized for each urotype.

There were no significant differences in the number of men with moderate to severe LUTS in each cluster, nor in the obstructive and irritative symptom scores.

Urotype	LUTS score (%), Moderate/severe	Score (median, [IQR])	
		Obstructive	Irritative
1 (n=85)	35 (41.2%)	1.0 [0.0, 5.0]	4.0 [2.0, 6.0]
2 (n=89)	43 (48.3%)	3.0 [1.0, 7.0]	4.0 [3.0, 7.0]
3 (n=65)	27 (41.5%)	1.0 [0.0, 4.0]	4.0 [3.0, 7.0]
4 (n=61)	24 (39.3%)	3.0 [0.0, 5.0]	3.0 [2.0, 6.0]
5 (n=59)	33 (55.9%)	3.0 [1.0, 6.0]	5.0 [3.0, 7.0]
6 (n=58)	27 (46.6%)	3.0 [0.0, 5.8]	3.0 [2.0, 6.0]
7 (n=37)	16 (43.2%)	2.0 [0.0, 5.0]	4.0 [2.0, 6.0]
8 (n=19)	4 (21.1%)	2.0 [0.0, 4.5]	3.0 [2.0, 5.0]
p-value	0.23	0.10	0.31

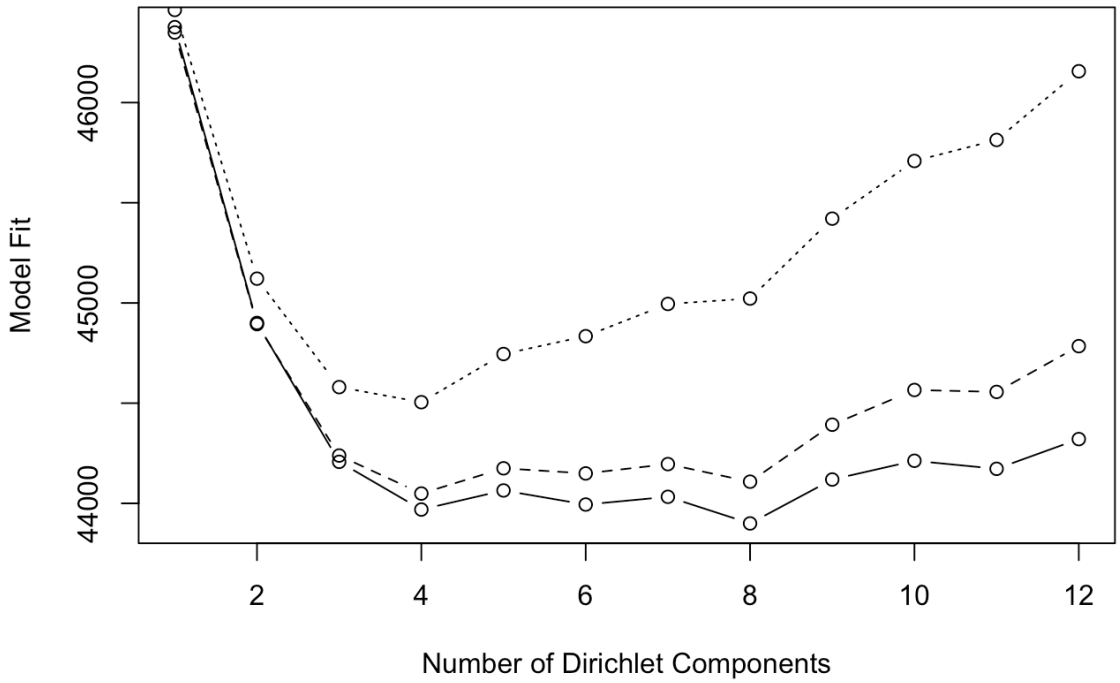


Figure 2.S6. Evaluation of the DMM model fit of varying numbers of clusters using the Laplace approximation, Aikaike Information Criterion (AIC), and Bayesian Information Criterion (BIC).

Smallest dotted line represents AIC, middle line is the BIC, and solid line is the Laplace approximation.

Chapter 3. Cell-free Microbial DNA in Blood is Associated with Prostate Cancer

Kate R. Bowie[1,2], Lisa Karstens[3,4]

1. Department of Biomedical Engineering, Oregon Health & Science University, Portland, Oregon, USA

2. Cancer Early Detection Advanced Research (CEDAR), Knight Cancer Institute, Oregon Health & Science University, Portland, OR, USA

3. Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, Oregon, USA

4. Department of Obstetrics and Gynecology, Oregon Health & Science University, Portland, Oregon, USA

3.1 Abstract

In the absence of infection, blood has previously been understood to be absent of proliferating microbes. Recent studies have challenged that dogma and researchers are now interrogating microbial DNA in the blood of both healthy and diseased individuals. Researchers have started to establish associations between circulating microbial DNA and prostate cancer, among other diseases. However the majority of these studies have relied on plasma. Here, we examine microbial DNA from the blood of 5 men without prostate cancer (PC), 5 men with low-grade PC, and 5 men with high-grade PC. To fully survey microbial DNA in blood we separated blood into fractions (plasma, red blood cells, and buffy coat) as well as added DNase to a subset of these fractions to determine if microbial-DNA is cell-free. We measured 16S rRNA gene (16S) copy number and submitted DNA samples for 16S rRNA gene sequencing. Our analyses revealed plasma had the least number of copies of 16S and least number of genera, while red blood cells had the most. We also established that the majority of microbial DNA in blood is cell-free and the composition of these samples is associated with PC. Our study demonstrates that microbial DNA can potentially be used as a biomarker in PC, however larger studies are needed to confirm our findings.

3.2 Introduction

Blood has long been considered sterile, or free of living and proliferating microbes. Recently, the potential existence of circulating microbes and microbial DNA in blood and its link to health and disease has garnered much interest. This is in part due to advances in sequencing technology allowing researchers to interrogate low-microbial biomass samples, like blood, more reliably.[135–137] Despite progress in sequencing and bioinformatic methods, microbes in blood remain controversial due to the lack of consensus on how best to handle low-microbial biomass samples and the high levels of contamination in such samples.[65] Although controversial, many believe the microbes found in circulation are part of a circulating microbiome,[138] while other researchers consider the microbes in circulation to be transitory.[78] Regardless, researchers are uncovering associations between circulating microbial DNA in blood and aspects of human health with potential clinical implications, such as being predictive of response to chemotherapy treatments.[95,139]

In relatively healthy populations, studies have shown up to 100% of samples having microbial DNA in blood while others determine as few as 16% of human blood samples have microbes.[78,79,140] This variation is likely due to differences in experimental approaches in sequencing, sample processing, and decontamination procedures. Although studies have pointed to the presence of *Staphylococcus* and *Cutibacterium*, there has not been agreement in specific bacteria found in the blood of healthy individuals indicating significant compositional heterogeneity.[78] When it comes to evaluating microbes and microbial DNA in blood in individuals with disease, researchers have identified associations between specific genera found in circulation and hypertension as well as type 2 Diabetes.[92,141] There have also been

associations between both abundance and diversity of the microbes in blood with liver cirrhosis, respiratory diseases, and chronic kidney disease.[82,84,91]

When it comes to cancer, approximately 16% of newly diagnosed cancers are caused by infectious agents, such as microbes.[142] In fact, bacteria have been found in tumors themselves,[59] and specifically implicated in prostate cancer (PC) as playing a role in inflammation which is important for the development of PC.[93] Importantly, Poore et al. found a distinct bacterial signature in the bloodstream of PC patients.[83] Additionally, microbial associations between the urinary microbiome and PC have been found.[76] To build on this work, we designed a study to rigorously evaluate microbial DNA from blood and urine as a potential biomarker for PC.

Overall, little is known about microbial DNA in blood, so we set out to uncover the localization of microbial DNA in blood. The majority of studies have relied on plasma as a proxy for blood, even though other components of blood also have microbial DNA.[79,83,94] Few studies have investigated other blood fractions (buffy coat and red blood cells), and none have investigated these fractions in disease. To build and improve on the knowledge of microbial DNA in blood, we measured the microbial load in each fraction treated with and without DNase, in non-cancer and prostate cancer patients. We also obtained matched urine samples from patients and performed a comparative analysis on bacteria found in blood and urine. Additionally, many studies have used 16S amplicon gene sequencing, which targets a small region of the 16S gene. We evaluated the microbial composition using synthetic long-read sequencing, which may improve signal to noise in low microbial biomass studies.

3.3 Results

Study Participants and Sequencing Overview:

Blood samples were taken from 15 men before undergoing prostate biopsy at Oregon Health & Science University. Age and PSA were recorded for all men (**Figure 3.1a**). The average age of participants was 65.8 years-old, with an average BMI of 28.8, and average PSA of 9.2 ng/mL. There were no significant differences in age, BMI, nor PSA found between disease status. The blood samples were processed producing plasma, red blood cell pellet (RBC), buffy coat layer (BF), and one aliquot of whole blood was retained (WB). Copies of 16S rRNA genes were measured with droplet-digital PCR (ddPCR) and the remaining sample was sent to Loop Genomics for synthetic full-length 16S rRNA gene sequencing (**Figure 3.1b**). All blood fraction samples generated sequences, with a median sequencing depth of 3684 reads (range 22-18,787 reads).

a

	Participants (n=15)	Screen negative (n = 5)	Low-grade PC (n = 5)	High-grade PC (n = 5)	p-value
Age (years), mean (s.d.)	65.8 ± 5.9	63.2 ± 6.9	65.8 ± 5.9	66.2 ± 5.6	0.15
BMI (kg/m²), mean (s.d.)	28.8 ± 4.22	29.6 ± 4.4	29.2 ± 4.5	27.7 ± 4.5	0.64
PSA (ng/mL), mean (s.d.)	9.2 ± 13.9	5.4 ± 3.4	9.2 ± 13.9	19.1 ± 21.7	0.47

b

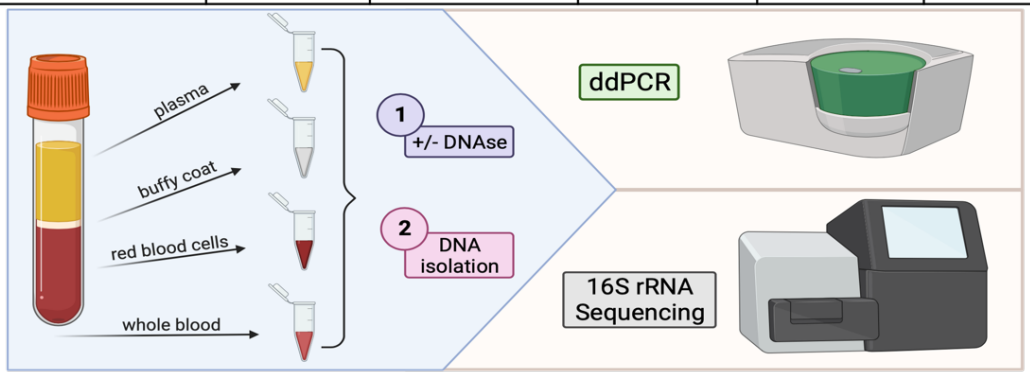


Figure 3.1. Experiment overview.

a) Table of clinical characteristics of patient cohort. There were no significant differences in age, BMI, nor PSA between groups. **b**) Cartoon depicting how blood was processed and the subsequent data generation.

DNase-Treated Blood Samples Indicate the Presence of Cell-Free Microbial DNA in Blood

First, we measured the copies of 16S rRNA genes in each sample. The samples were normalized to volume of whole blood used to produce each fraction. Surprisingly, BF had the lowest copies of 16S with a mean of 23.3 copies/µL (interquartile range [IQR] 15.8, 62.8 copies/µL) with plasma having slightly more (median 32.5, IQR: 16.7, 38.8 copies/µL, **Figure 3.2a**). RBCs followed (median 55.9, IQR: 14.7, 80.4 copies/µL), and as expected, WB had the highest copies of 16S (median 167.9, IQR: 132.0, 209.8 copies/µL). The amount of 16S genes in each fraction was significantly different (p-value = 6.095e-06).

To identify if the microbial DNA in each fraction was intracellular or extracellular, we compared each fraction with and without DNase added (**Figure 3.2b**). The DNase should remove extracellular DNA and leave only the intracellular microbial DNA to be measured by ddPCR. As plasma contains only cell-free DNA, it should have no copies of 16S genes when DNase has been added. As expected, the DNase-treated plasma samples had a median of 0.0 copies/µL of 16S (**Figure 3.2b**). The RBC DNase samples had the second lowest number of 16S copies (median 1.2 copies/µL, IQR: 0.0, 14.2 copies/µL). After plasma and RBC, the BF DNase samples had a median of 12.8 copies/µL, and lastly WB had the most copies of 16S after DNase treatment with a median of 44.0 copies/µL of 16S (IQR: 8.8, 165.9 copies/µL). All the DNase samples had significantly less copies of 16S than their non-DNase counterparts (**Figure 3.2b-c**), which indicates the presence of cell-free microbial DNA in every blood fraction.

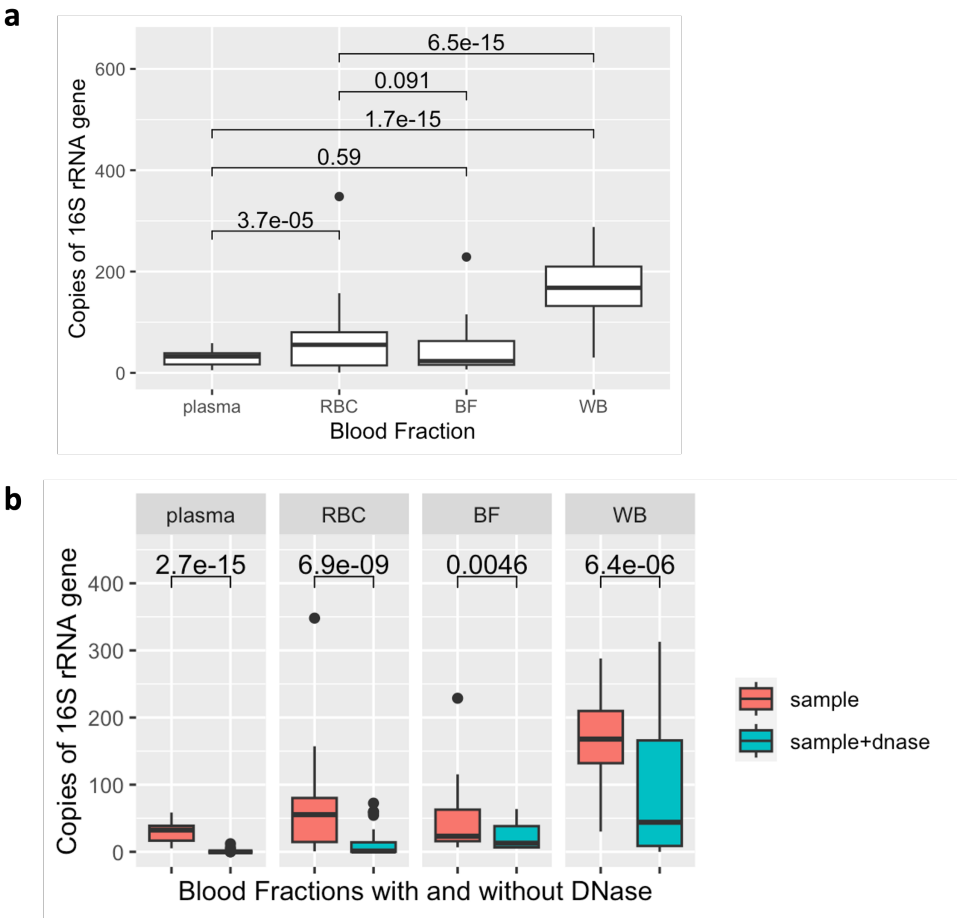


Figure 3.2. Evaluation of microbial load in blood.

a) Boxplot displaying the number of copies of 16S rRNA genes per µL across blood fraction types. The blood fractions had significantly different number of copies determined by Kruskal-Wallis test (p-value = 6.095e-06). **b**) Boxplot comparing each fraction to its associated DNase-treated counterpart. All fractions had significant differences in copies/µL of 16S when compared to DNase samples.

Negative Controls Look Distinct from Blood Fractions

After establishing the presence and quantity of microbial DNA in the blood fractions, we evaluated the composition of the samples using 16S rRNA gene sequencing. We first examined the positive and negative controls to assess confidence in these low-microbial biomass sample sequencing results. The positive controls were mock microbial dilution series sequenced along with the blood fraction samples (**Figure 3.S1**). Sequencing revealed all eight expected genera in the positive controls regardless of dilution. The average percentage of unexpected taxa detected in the positive controls was less than 0.5%.

Negative controls were microbial-free water blanks that were processed alongside the blood fraction samples. The majority (87.0%) of the negative controls generated sequences with an average of 1,121 reads (minimum 1 read, maximum 7,693 reads). The five most abundant genera in our negative controls were *Bacillus*, *Legionella*, *Prevotellaceae*, *Prauserella*, and *Escherichia-Shigella*, of which *Bacillus*, *Legionella* and *Escherichia* have previously been described as typical contaminating sequences.[143] The top five most abundant species in the negative controls were *Lactobacillus fermentum*, *Bacillus mycoides*, *Rubrobacter bracarensis*, *Pseudomonas lutea*, and *Methylobacterium adhaesivum*. To determine if the genera associated with the negative controls should be removed, we examined how prevalent these genera were in the associated fraction (**Figure 3.3a**). Outside of *Prauserella* and *Rubrobacter*, we did not identify substantial overlap between the taxa found in the samples and those of the negative controls. *Prauserella* was present in 35.0% of samples and in 53.8% of negatives (**Figure 3.3b**). As *Prauserella* has only been reported to be found in lakes and sediment,[144] we removed all sequences mapping to *Prauserella*. Similarly, sequences mapping to *Rubrobacter* were present in both negative controls and samples and were removed from the data. Nitrosomonas, Rhodospirillum,

Rhodobacter, and Salinispora were also removed as they have not been previously found to be associated with the human body.[145–147]

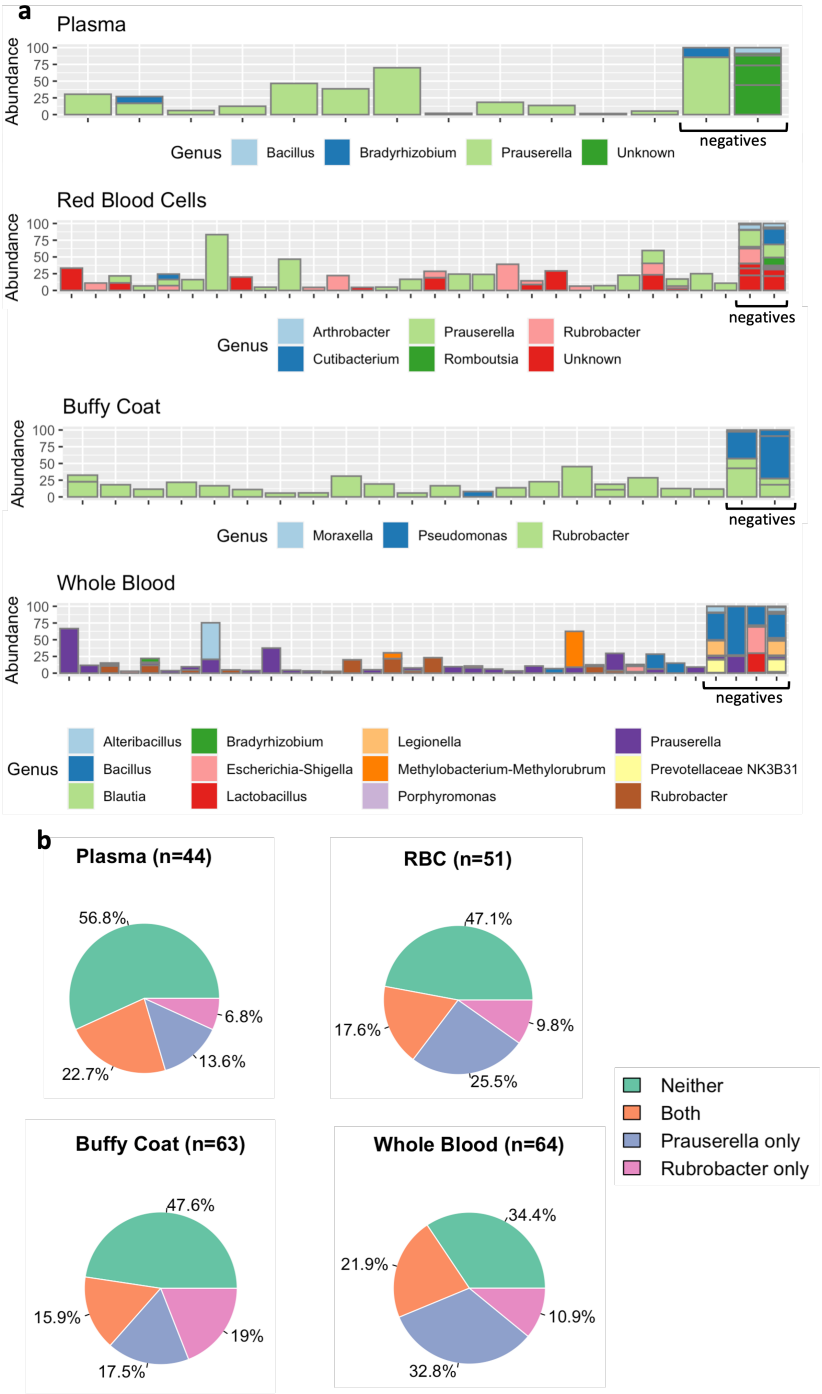


Figure 3.3 Bacteria from negative controls across samples.

a) Stacked bar plots of the abundance of genera for each sample. All genera selected for these plots were based on the negative controls for the associated blood fraction. **b**) Pie charts demonstrating that the majority of samples had *Prauserella*, *Rubrobacter*, or both – which are considered contaminant genera in this dataset.

Diversity of Microbial DNA Varies by Blood Fraction

After removal of contaminants, we found there were significantly different numbers of reads for each fraction ($p = 0.0005$, **Figure 3.S2a**), which was similar to the ddPCR results. Plasma had the least number of reads (median 1313 reads), followed by RBC, WB and finally BF (**Figure 3.S2b**, RBC: median 3391; WB: median 6807, and BF: median 7166;). All samples were rarefied to 50 reads, which removed 13.3% of plasma and BF samples, 40.0% of RBC samples, and 53.3% of WB samples. No association was found with health status and successfully sequenced samples, nor sampling depth, regardless of fraction. The sequencing data revealed 12 phylum, 103 genera and 159 species present in the blood samples. Compositionally, Firmicutes was the most abundant phylum followed by Proteobacteria, Actinobacteria, and Bacteroidota (**Figure 3.4a**). The most abundant genera for all fractions were *Escherichia-Shigella, Bacillus, Bifidobacterium, Pseudomonas,* and *Staphylococcus*. All have previous been found in the human body.[148] In terms of individual fractions, there was significant overlap of the top five genera, shown in **Figure 3.4b**. There were however no significant differences in the relative abundances of the top 10 genera between the fractions. The most abundant species were *Bifidobacterium adolescentis, Escherichia coli, Bacteroides thetaiotaomicron, Bacillus subtilis*, and *Staphylococcus aureus*. Interestingly, only 47.8% of the taxa were assigned at the species level, and the rest could not be assigned, hence we continue the remaining analysis at the genus level.

Many of the studies examining microbial DNA in blood use plasma or serum, therefore we were curious if the alpha diversity of plasma was similar to the other fractions. All three of the alpha diversity metrics (observed, shannon index,

and inverse simpson index) were significantly different between blood fractions, with plasma having the least observed genera (mean 7.4 genera, range 2-11 genera) and RBC having the most of the blood fractions (mean 10.8 genera, range 1-15 genera), **Figure 3.4c**). Next, we were interested in whether fractions are more similar to each other or to the patient they came from, so we performed beta diversity analysis. However, we did not find any significant associations between any beta diversity metric (weighted UNIFRAC, unweighted UNIFRAC, and Bray-Curtis) and patient nor fraction (**Table 3.S1**).

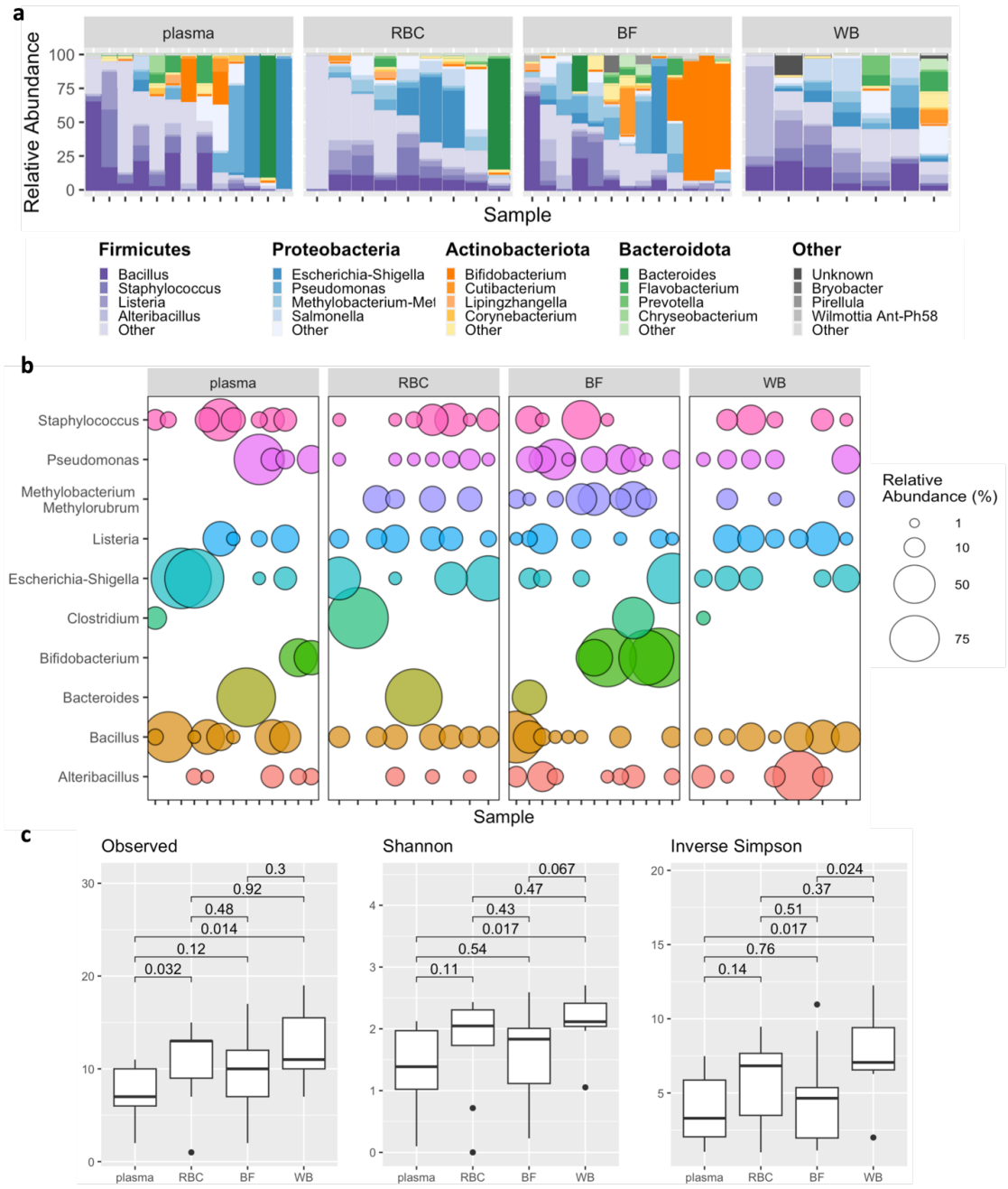


Figure 3.4. The blood fractions share similar genera; however their alpha diversity is significantly different.

a) Stacked bar plot of the microbial DNA found in all blood fractions ordered by Firmicutes abundance. **b**) Bubble plot of the top 10 genera separated by blood fraction. The fractions share the majority of the same genera and the bubbles display the relative abundance of each genus. **c**) Boxplots of the three alpha diversity metrics (observed, shannon, and inverse Simpson indices) demonstrating plasma has the lowest alpha diversity of all the fractions.

Composition of DNase-Treated Samples Reveal Majority of Microbial DNA in Blood is Cell-Free

Little is known about where microbial DNA in blood is located. We sought to understand if there are compositional differences between intracellular versus extracellular microbial DNA. Only 54.9% of the DNase samples were successfully sequenced (**Figure 3.S3**), whereas 100% of samples without DNase successfully sequenced. The DNase samples had significantly less reads than those without DNase ($p = 0.001$, **Figure 3.5a**), which is expected as the samples had less overall microbial DNA (**Figure 3.3a**). By number of reads, the DNase samples seemed to be largely dominated by a few genera (**Figure 3.5c**), with the top five being *Clostridium, Bacteroides, Escherichia-Shigella, Bacillus,* and *Bifidobacterium.* Interestingly, Clostridium had higher relative abundance in DNase treated samples than samples without DNase, suggesting a potential mechanism for intracellular translocation (**Figure 3.5d**). We next measured alpha diversity and identified a significant difference between the diversity of samples and DNase samples (**Figure 3.5b**). DNase samples had a median of 3 observed genera (range 1-16) while the samples without DNase had a median of 10 observed genera (range 1-19) – a trend that held true for the Shannon and inverse Simpson indices (**Figure 3.5b**).

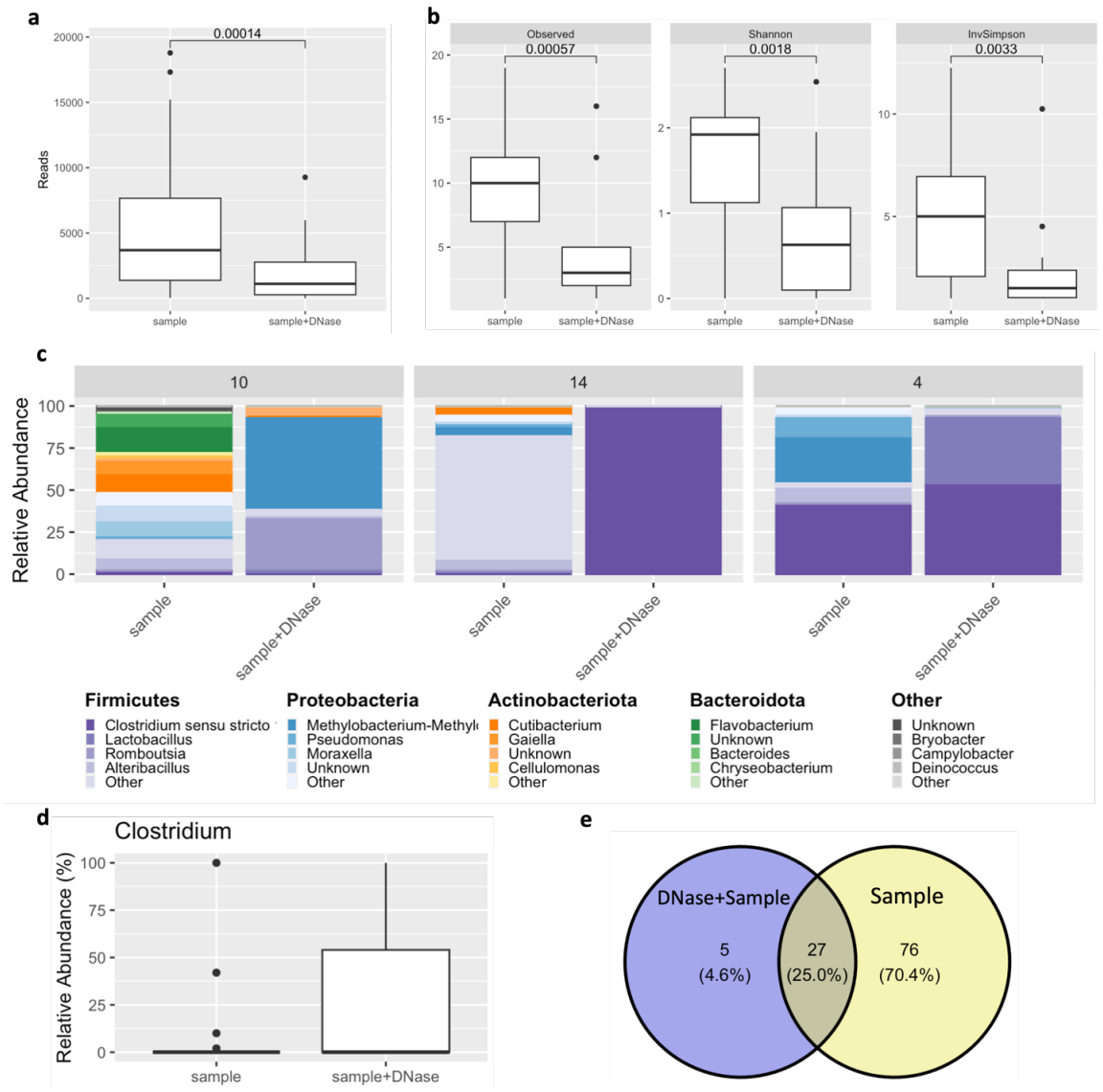


Figure 3.5. DNase blood samples have less diversity than non-DNase samples.

a) Boxplot displaying the significant difference in sequencing reads between DNase and non-DNase samples (p= 0.0001). **b**) Summary of alpha diversity metrics between DNase and non-DNase samples, demonstrating the significant decrease in alpha diversity in DNase samples. **c**) Stacked bar plot of microbial DNA found in paired DNase and non-DNase blood fractions for patients 10, 14, and 4. **d**) Boxplot showing that DNase samples have higher relative abundance of *Clostridium* than non-DNase samples. **e**) Venn diagram indicating the number of distinct and shared taxa between DNase and non-DNase samples.

We evaluated beta diversity metrics of samples with and without DNase and found significant associations with both weighted UNIFRAC and the Bray-Curtis beta diversity metrics (**Figure 3.S3b-c**). We also identified 76 genera found only in the non-DNase samples, meaning these genera are only located extracellularly in this patient population (**Figure 3.5e**). Of those 76, 15 were not assigned at the Genus level. Conversely, we identified 5 genera present only in the DNase samples. This could be due to the DNase samples having less diversity, which potentially allows for more rare genera in these samples to be resolved.

Prostate Cancer-Associated Patterns in Beta Diversity of Microbial DNA in Blood

Next, we wanted to examine if there are any disease-associated patterns in microbial DNA in blood. Our patient population consists of 15 men, made up of three groups of five men each. The groups consist of individuals who are screen-negative for PC, low-grade PC, and high-grade PC. Each individual should have four samples – one for each blood fraction – for a total of 20 samples per disease group. After sequencing, we were left with 13 screen-negative samples, 18 low-grade PC samples, and 14 high-grade PC samples (**Figure 3.6a**). We evaluated alpha diversity between disease groups and did not find any associations (**Figure 3.S4c**). We did however identify significant associations between all three beta diversity metrics (weighted UNIFRAC: $p = 0.01$, unweighted UNIFRAC: $p = 0.003$, and Bray-Curtis: $p = 0.04$) and disease type (**Figure 3.6b, Figure 3.S4a-b**). This indicates that there are compositional differences in the microbial DNA between men that are screen-negative, diagnosed with low-grade PC, and diagnosed with high-grade PC. Next, we

created a biplot at the phylum level to begin to uncover which taxa could be responsible for the compositional difference between disease states (**Figure 3.6c**). This biplot, along with an overview of the data (**Figure 3.6d**) demonstrated that many of the high-grade PC samples were missing Bacteroidota, and it could potentially be a marker of more severe disease. However, the differences in relative abundance of Bacteroidota across all fractions was a trend and not significantly different ($p = 0.07$, Kruskal-Wallis test, **Figure 3.6d-e**). We did discover an association with the relative abundance of Proteobacteria of all fractions combined between disease types ($p = 0.02$, **Figure 3.6f**), with high-grade PC having the highest abundance. We also examined associations between alpha diversity and PSA, BMI, and age, but did not find any significant associations (**Figure 3.S4**).

a

Samples after sequencing	Screen-negative	Low-grade PC	High-grade PC
plasma	4	4	5
RBC	3	3	3
BF	4	4	5
WB	2	3	2

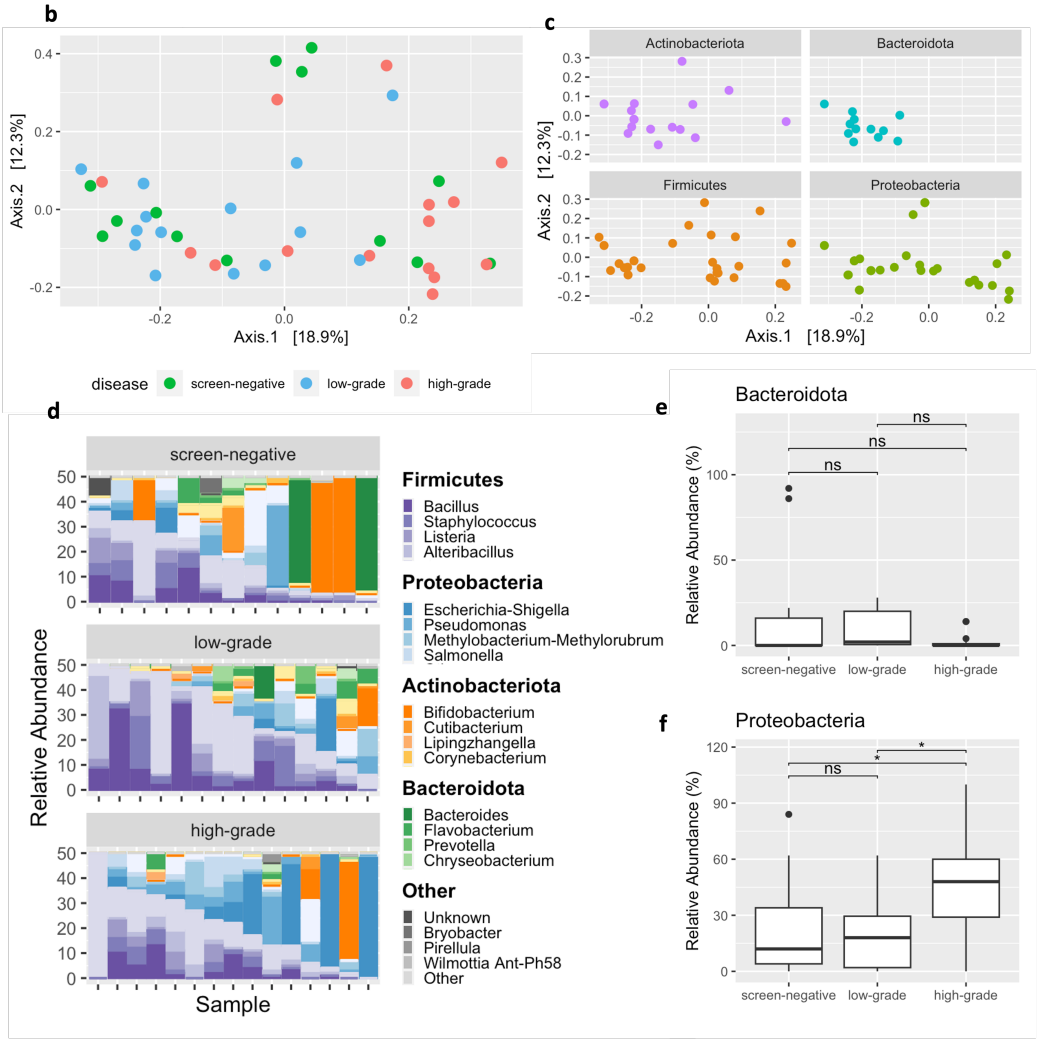


Figure 3.6. Beta diversity analysis reveals disease associated patterns in microbial DNA in blood.

a) Table of the final number of blood fraction samples in each disease group used for analysis. **b**) Unweighted UNIFRAC principal coordinate analysis (PCoA) plots of all blood fractions colored by disease status (p=0.003). **c**) Biplot demonstrating the phyla associated with each sample's composition. **d**) Stacked bar plots faceted by disease for all fractions. **e**) Boxplots showing the relative abundance of Bacteroidota and **f**) Proteobacteria for all disease status.

Microbial DNA in Urine is Distinct from Blood

The prostate is located right underneath the bladder, thus we decided to investigate the microbiome in urine samples from the men in our cohort. Of the cohort, three urine samples were removed due to insufficient sequencing reads (<50). We were interested in the overlap between the microbial DNA found in blood and urine samples, as well as if there are disease-associated patterns in urine. We found interestingly that nine of the 12 urine samples had overlap in genera present with its corresponding blood samples for individuals in our cohort (**Table 3.S2**). The genera overlap between blood and urine samples ranged from 0.0% to 25.0% of the individual's total genera, with a mean overlap of 6.8%. *Bacillus*, *Prevotella*, and *Staphylococcus* each had overlap in three individuals. Despite the overlap, we found that the composition of urine is distinct from the blood samples as evaluated by beta diversity (**Table 3.S3**).

3.4 Discussion

The results from our study indicate that there is measurable microbial DNA in circulation. We identified 103 genera and 159 species, however not all taxa had taxonomic assignment. If we look at assigned taxa, we identified 85 genera and 76 species. In a similar study, Tan et al found 56 genera and 117 species. However, this study had a much larger number of participants (10,000) and used a different sequencing modality and decontamination parameters which could explain the

disparity in number of taxa.[78] This discrepancy again points to the importance of sequencing and data processing and how they can impact the resulting data.

We found that BF contains the least amount of microbial DNA (**Figure 3.2a**), although plasma had a lower mean (28.4 copies/µL versus BF at 41.5 copies/µL) and had the fewest number of genera (**Figure 3.4c**), indicating a potential limitation of prior work that relied only on plasma. Our results also demonstrated that the RBC fraction has more copies than BF (**Figure 3.2a**). These results are contrary to a previous study, although we utilized ddPCR which is known to produce more precise and reproducible results as compared to qPCR.[79,149] We had expected BF to have the highest copy number and number of genera since white blood cells as part of the immune system are understood to engage with microbes, so we had assumed there would be more associated microbial DNA. Unexpectedly, we also discovered the RBC fraction to have on average more observed genera than BF (median 13 genera compared to 10 genera, **Figure 3.4c**) – a finding that is similar to the blood fraction study by Paisse et al.[79] These findings suggest an interaction between RBCs and circulating or transitory microbes. In fact, in conditions of bacteriemia, RBCs have been documented interacting with bacteria by either entrapping, killing, then releasing the dead bacterium back into circulation,[150] or the RBCs use their electric charge to attract and kill bacteria.[151] Our findings could point to similar mechanisms of RBCs as seen in bacteremia. Not only this, but RBCs have a life cycle of approximately 120 days, thus studying the RBC fraction for microbial DNA could serve as a recent history of the types of microbes that have translocated in the body.[152]

A major finding of our study is that the majority of microbial DNA in blood is cell-free. No other studies have used DNase to explore the question of intra- versus extracellular microbial DNA. We found the DNase-treated samples to have

significantly less copies of 16S per µL (**Figure 3.2b-c**) and complexity (**Figure 3.5b-c**). We did however discover the DNase-treated samples shared 84.4% of their genera with the non-DNase treated samples (**Figure 3.5e**), and had higher relative abundance of *Clostridium*. We used NCBI Blast and found the *Clostridium* sequence to map to *Clostridium sporogenes* – which is a spore-forming bacteria and while uncommon, has previously been documented to cause bacteremia.[153] *C. sporogenes* is a member of the normal gut microbiome, however in the spore state, *C. sporogenes* are dormant and can survive in hostile environments even in the absence of nutrients,[154] perhaps allowing it to go undetected in blood. Many of our DNase-treated samples did not successfully sequence, which inhibited the analysis we could perform. Future studies in this area should include examining how disease potentially affects the composition in DNase-treated samples, which could indicate changes to the cells that make up the RBC and BF fractions.

Supporting prior research of circulating microbial DNA associations with disease, we did determine compositional differences between PC-negative samples as well as those with low- and high-grade disease. With a larger sample size, a more robust analysis of the alpha diversity differences in specific fractions with respect to disease could give us insight into how disease impacts these transitory microbes. Specifically, more samples are needed to study the potential trend of *Bacteroidota* missing from high-grade PC samples regardless of fraction (**Figure 3.6b-c**). What is unclear is if the patterns we discovered are a general cancer-related finding, or specific to PC. Regardless, we believe this supports previous work demonstrating that microbial DNA in blood has disease associations.[82–84,92,141]

A limitation of this study was the small sample size, which was a trade off we made to increase rigor. As we processed every fraction in triplicate and doubled that to treat samples with DNase, we processed a total of 360 samples not including the

controls. We believe the replicates are important for capturing the diversity of these low-microbial biomass samples, however a small study of 15 patients quickly became cumbersome. We also believe in the future alternative taxonomic classification methods should be explored achieving species level assignments for long read sequencing.

In summary, our study indicates the presence of circulating microbial DNA in both PC negative and positive men. Our results suggest microbial DNA in blood is sparse, heterogenous, and present in all fractions of blood. We revealed that not only is the majority of microbial DNA in blood cell-free, but also that the majority of the microbial diversity stems from the RBC fraction (as opposed to plasma or BF fraction). More studies with larger sample sizes are needed to further investigate these associations with specific fractions of blood and disease.

3.5 Methods

Patient Cohort

Our research complies with all relevant ethical regulations of Knight Cancer Institute at Oregon Health & Science University. Patients were undergoing prostate biopsy and were consented through VAPORHCS/OHSU: Cancer Early Detection Advanced Research (CEDAR) Specimen and Data Repository. The associated IRB# is 18048 at Oregon Health & Science University, and IRB# 4214 at Portland VA Medican Center. Participants were not compensated.

DNA Isolation and Sequencing

Blood was drawn from men with high-grade PC, low-grade PC, and without cancer (n=5 per group). Samples were processed into the plasma, buffy coat, red blood cell pellet (RBC) and whole blood fractions in six 200 µL aliquot per fraction prior to freezing. For each sample, 5 mL of blood was removed from the tube and 600 µL whole blood was removed, followed by a room temperature 1200 rpm centrifugation step for the plasma fraction. The remaining 8-10 mL underwent density gradient centrifugation using Ficoll-Paque PLUS (Cytiva) for the buffy coat and red blood cell pellet fractions. To examine if microbial DNA was extracellularly or intracellularly located, DNase (New England Biosciences) was added to three of the six replicates for each fraction, and the product protocol used. Once all blood fractions were aliquoted into three replicates and DNase added to the remaining three, samples were frozen. Microbial-free water (Qiagen) was used as negative controls and a mock microbial community (Zymo) was serially diluted per previous research,[65] and processed alongside blood samples. Microbial DNA was extracted using the QiAMP DNA mini kit (Qiagen). Extracted DNA from the blood fractions was submitted to Element Biosciences for their 16S rRNA synthetic long-read sequencing using LoopSeq technology.

Droplet-Digital PCR

Droplet-digital PCR (ddPCR) was used to measure 16S rRNA copies. Each PCR reaction was prepared in a PCR hood in a dedicated ddPCR room. For the mock bacterial community dilution series, each dilution was diluted either 1:1000 or 1:100 in microbial-free water (Qiagen). The QX200 Digital PCR System with Auto Droplet Generator and Reader along with the QX200 ddPCR EvaGreen Supermix

(Biorad) was used. The PCR protocol uses a 2˚C/second ramp rate and starts with a 10-minute 95˚C enzyme activation step followed by 40 cycles of a two-step protocol (94˚C for 30 seconds and 59˚C for 1 minute), and lastly cycles up to 98˚C for 10 minutes. V6 primers from the Sfanos lab at Hopkins were used (Forward: CAACGCGWRGAACCTTACC; Reverse: CRRCACGAGCTGACGAC). The results were then normalized to starting whole blood volume required to produce 200 µL of each blood fraction for comparison across fractions.

Bioinformatics and Statistical Analyses

Loop Genomics provided us with assembled full-length 16S rRNA gene sequences. These raw sequences were processed into amplicon sequence variants (ASVs) using DADA2.[137] The RDP Classifier was used to map the ASVs to the SILVA 138 16S rRNA reference set for taxonomic identification. All analyses were completed on sequences between 1400 and 1550 basepairs as per LoopSeq protocol (Element Biosciences). The mock community dilution series and negative controls were analyzed to determine how to handle biological samples using phyloseq (version 1.42.0) and visualized using microshades (version 1.11).[134] All subsequent analyses were done on ASVs agglomerated at the Species or Genus level in R. The replicates for each sequencing sample were combined prior to rarefaction to represent one microbiome composition per fraction per individual. Rarefaction to 50 reads was based on retaining the maximum number of samples while accurately representing the composition (**Figure 3.S5**). The *Vegan* R package version 2.6.4 and *rstatix* version 0.7.2 were used for all statistical analyses. We also looked at relative abundance of the top 5 phyla and top 10 genera. A Kruskal-Wallis test was used to test for a

significant association between relative abundance and clinical characteristic of interest, and then a pairwise Wilcoxon Rank Sum test with FDR correction.

3.6 Supplemental Figures and Tables

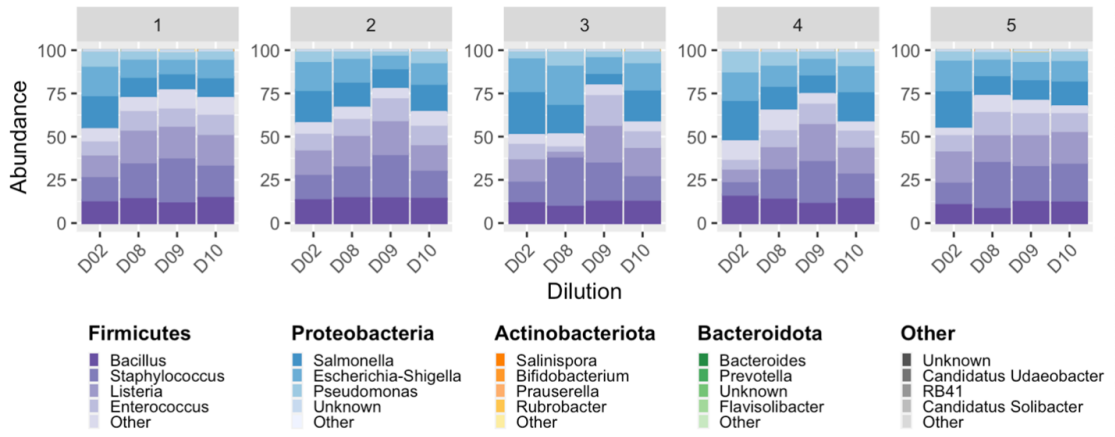
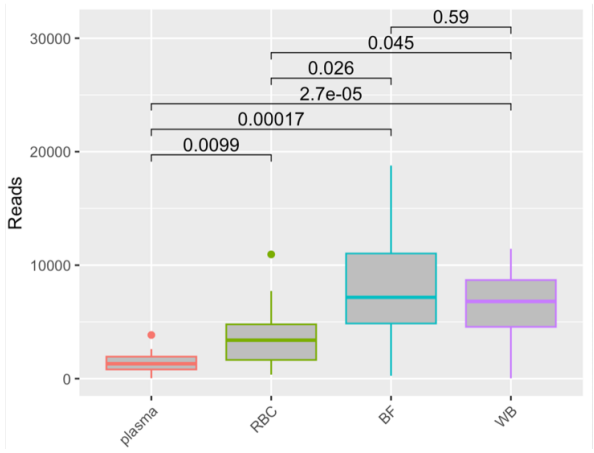


Figure 3.S1. Stacked bar plot of mock community dilution series faceted by extraction batch.

The samples were separated into six different batches for DNA extraction. Each extraction batch and four mock community dilutions, with D02 being the most concentrated and D10 the most dilute. All eight expected taxa were present in every dilution and batch.

a



b

Fraction	Reads (mean)	Range (min, max)
plasma	1450.5	40, 3847
RBC	3889.0	356, 10951
BF	7943.1	252, 18787
WB	6449.3	32, 11445

Figure 3.S2. Overview of sequencing reads by blood fraction.
a) Box plot of sequencing reads for each fraction (n=15). **b)** Table with a summary of the mean and range of sequencing reads for each fraction.

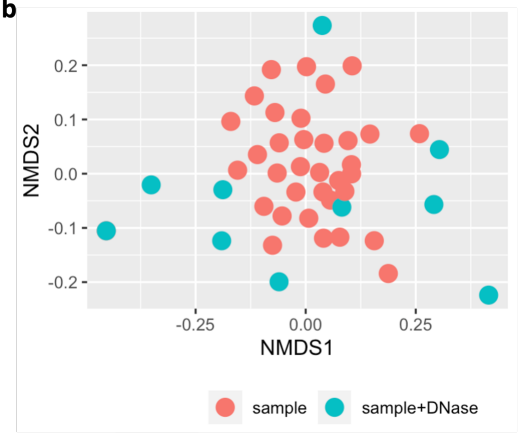
Table 3.S1. Summary of the p-values of beta diversity metrics.
The microbial composition of the plasma, RBC, and BF samples were tested for associations with fraction and patient. No significant associations were detected.

	Fraction	Patient
Weighted UNIFRAC	0.09	0.44
Unweighted UNIFRAC	0.30	0.99
Bray-Curtis	0.14	0.22

a

Fraction	# Samples+DNase Submitted	# Samples+DNase Sequenced	# Samples+DNase for analysis (%)
Plasma	2	2	1 (50.0%)
RBC	11	7	4 (36.4%)
BF	21	10	5 (23.8%)
WB	17	9	3 (17.6%)

b



c

Fraction	DNase	Fraction	Patient
Weighted UNIFRAC	**0.05**	0.46	0.30
Unweighted UNIFRAC	0.08	0.44	0.89
Bray-curtis	**0.04**	0.49	0.19

Figure 3.S3. DNase-treated samples had significantly different compositions than untreated samples.

a) Table showing that the majority of DNase-treated samples were not sequenced successfully. **b)** NMDS plot of the bray-curtis index demonstrating the compositional differences between DNase-treated samples and untreated samples. **c)** Table of the beta diversity metrics showing significant associations between the weighted UNIFRAC and bray-curtis metrics and DNase-treated and untreated samples.

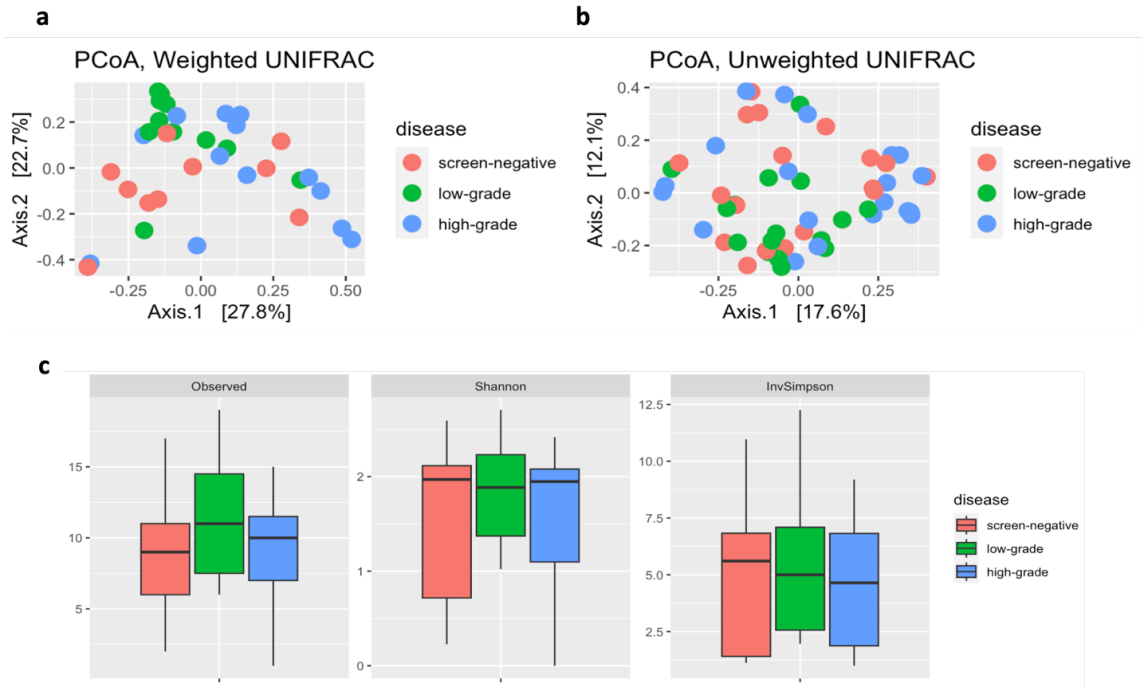


Figure 3.S4. Diversity plots of microbial DNA in blood fractions by disease status.

a) Weighted UNIFRAC (p=0.01) and **b)** unweighted UNIFRAC principal coordinate analysis (PCoA) plots of all blood fractions colored by disease status (p=0.003). **c)** Box plots showing no difference in alpha diversity between disease status for all four alpha diversity indices.

Table 3.S2. Summary of the genera unique to blood, urine, or shared between both specimens for each individual.

Patient ID	Unique in Blood	Unique in Urine	Shared
2	5	1	2 (Lactobacillus, Escherichia)
3	23	6	2 (Prevotella, Streptococcus)
4	13	12	0
5	18	6	4 (Prevotella, Veillonella, Campylobacter, Porphyromonas)
6	13	36	2 (Bacillus, Clostridium)
9	21	8	3 (Negativicoccus, Aerococcus, Staphylococcus)
11	19	2	1 (Staphylococcus)
12	13	3	0
13	30	1	5 (Bacillus, Listeria, Enterococcus, Lactobacillus, Staphylococcus)
14	30	5	1 (Prevotella)
15	23	39	1 (Bacillus)
16	13	6	0

Table 3.S3. Summary of the p-values for beta diversity analysis including both urine and blood samples from our patient cohort.

	Fraction	**Patient**
Weighted UNIFRAC	0.003	0.83
Unweighted UNIFRAC	0.001	0.97
Bray-Curtis	0.001	0.63

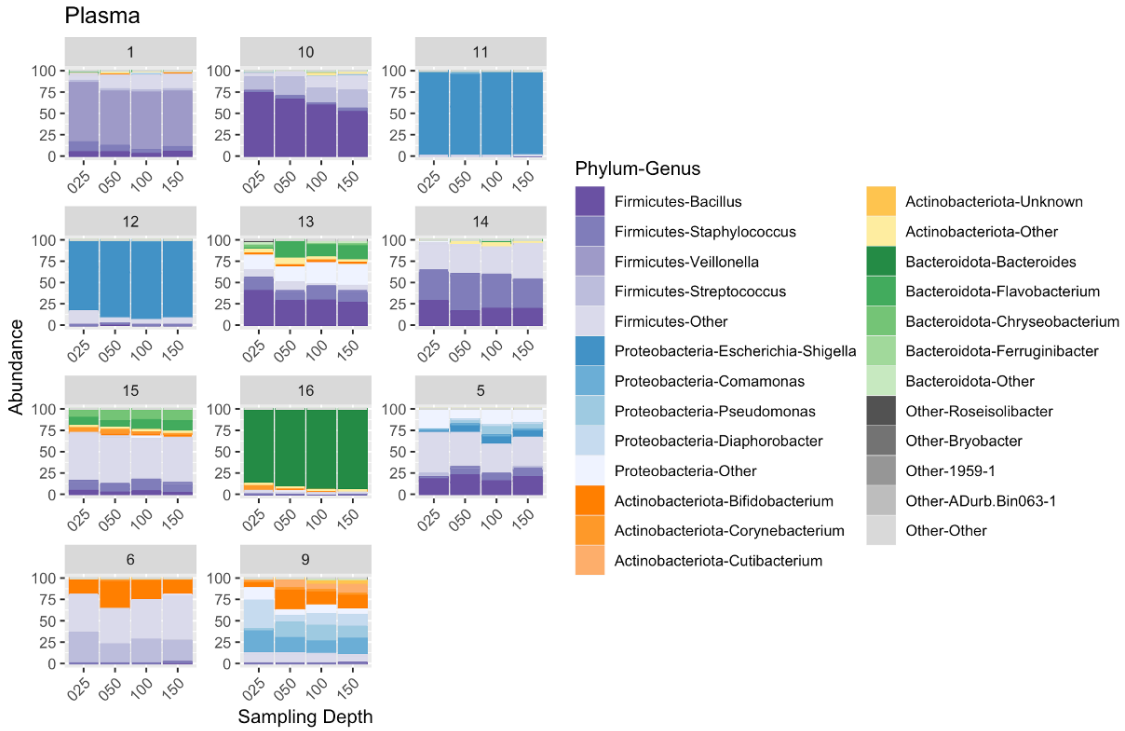


Figure 3.S5. Stacked bar plots of plasma samples rarefied to 25, 50, 100, and 150 reads.

Similar patterns were seen in the remaining fractions. Small compositional difference were seen between 50 and 150 reads, which justified using 50 reads for rarefaction to retain more samples.

Chapter 4. To Flow or not to Flow: Considerations for Quantitative Microbiome Profiling

Kate Bowie[1,2], Alec Furst[3], Fatoumata Jallow[4], Sean Davin[3], Lisa Karstens[5,6]

1. Department of Biomedical Engineering, Oregon Health & Science University, Portland, Oregon, USA

2. Cancer Early Detection Advanced Research (CEDAR), Knight Cancer Institute, Oregon Health & Science University, Portland, OR, USA

3. Division of Arthritis and Rheumatic Diseases, Oregon Health & Science University, Portland, Oregon, USA

4. Portland State University, Portland, OR, USA

5. Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, Oregon, USA

6. Department of Obstetrics and Gynecology, Oregon Health & Science University, Portland, Oregon, USA

4.1 Abstract

16S rRNA gene sequencing is often used to study bacteria in mixed communities. To measure the abundance of specific bacteria more accurately, many researchers have started combining quantitative methods with sequencing data, called quantitative microbiome profiling (QMP). However, there is little guidance on how proposed QMP methods perform. We have benchmarked a QMP method from Ott et al. by evaluating flow cytometry, droplet-digital PCR, and qPCR with 16S rRNA gene sequencing data of an eight-fold dilution series of a mock microbial community.[155] We established that flow cytometry and ddPCR were highly accurate and precise, while qPCR was not accurate nor precise, but had low variability. When combined with sequencing, we determined that qPCR should not be used with this QMP method with high-microbial biomass samples, and flow and ddPCR are appropriate for low-microbial biomass samples.

4.2 Introduction

Culture-independent methods are changing the way scientists interrogate microbes in and on the human body. Specifically, targeted sequencing of a conserved bacterial gene, the 16S rRNA gene (16S), has led to growth in the microbiome field as researchers can now inexpensively study microbes with unknown culture conditions. Sequencing of the 16S rRNA gene also proves useful in samples with high amounts of host DNA contamination since it only targets prokaryotes. 16S sequencing is both inexpensive and high throughput, generating massive amounts of microbiome compositional information. Many researchers also use 16S sequencing to determine relative abundances of taxa in a community, which is the proportion of sequencing reads for a specific taxon relative to all taxa present in the sample. However relative abundance has questionable biological relevance.[156]

Although advances in sequencing technologies have allowed for cheaper and more high-throughput data creation, data analysis methods have not advanced at the same rate. Majority of microbiome publications focus only on the most abundant microbes when evaluating either 16S or whole genome sequencing data.[157] 16S sequencing has led to improvements in our understanding of the host-microbe interactions, however it has many limitations. These limitations include ill-defined biased towards certain taxa during both DNA extraction and amplification, as well as throughout sequencing.[156,158] For example, bacteria differ in how easily they are able to by lysed as well as in how many copies of 16S they have, which can skew sequencing results.[159] To overcome these limitation, quantitative microbiome profiling (QMP) – methods which combine sequencing and quantification of 16S copies or bacterial cells – have been proposed.[155,160]

While QMP is a desirable method to obtain more relevant data about the microbiome, incorporating measurements of the microbiota needs to be done so reliably. Recent QMP methods include correcting sequencing reads by 16S gene copy number and subsampling sequencing reads based on quantity of bacterial cells.[155,160] As QMP is gaining popularity, we have benchmarked one method of combining sequencing and quantitative data proposed by Ott, et al. (2021).[155] We use samples with a known composition and in known quantities to investigate if this QMP method is appropriate for both high- and low-microbial biomass samples, as well as the limitations of the methods. We have compared three quantitative modalities: quantitative PCR (qPCR), droplet-digital PCR (ddPCR), and flow cytometry (flow), and evaluated each method on precision, accuracy, and variability as modeled on a mock bacterial community dilution series. We also combined each of the three modalities independently with mock microbial dilution series sequencing data and evaluated how well they performed.

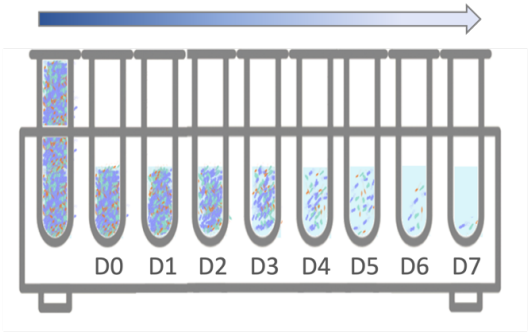
4.3 Results

Overview of Mock Microbial Dilution Series

To understand how well QMP works on both high- and low-microbial biomass samples, we serially diluted a commercially available ZymoBIOMICS Microbial Community Standard. This mock community underwent eight rounds of serial 3-fold dilutions with dilution zero (D0) being the undiluted and most concentrated sample, and dilution 7 (D7) as the most dilute sample (**Figure 4.1a**). The expected concentration of the mock community ranges from D7 (1.60×10^5 cells per mL or 4.64×10^9 copies of 16S per mL), to D0 (1.05×10^{10} cells per mL or 3.05×10^5 copies of

16S per mL). The abundance of the eight genera of the mock community vary depending on measurement by genomic DNA, cell number, or copies of 16S rRNA gene (**Figure 4.1b**).

a



b

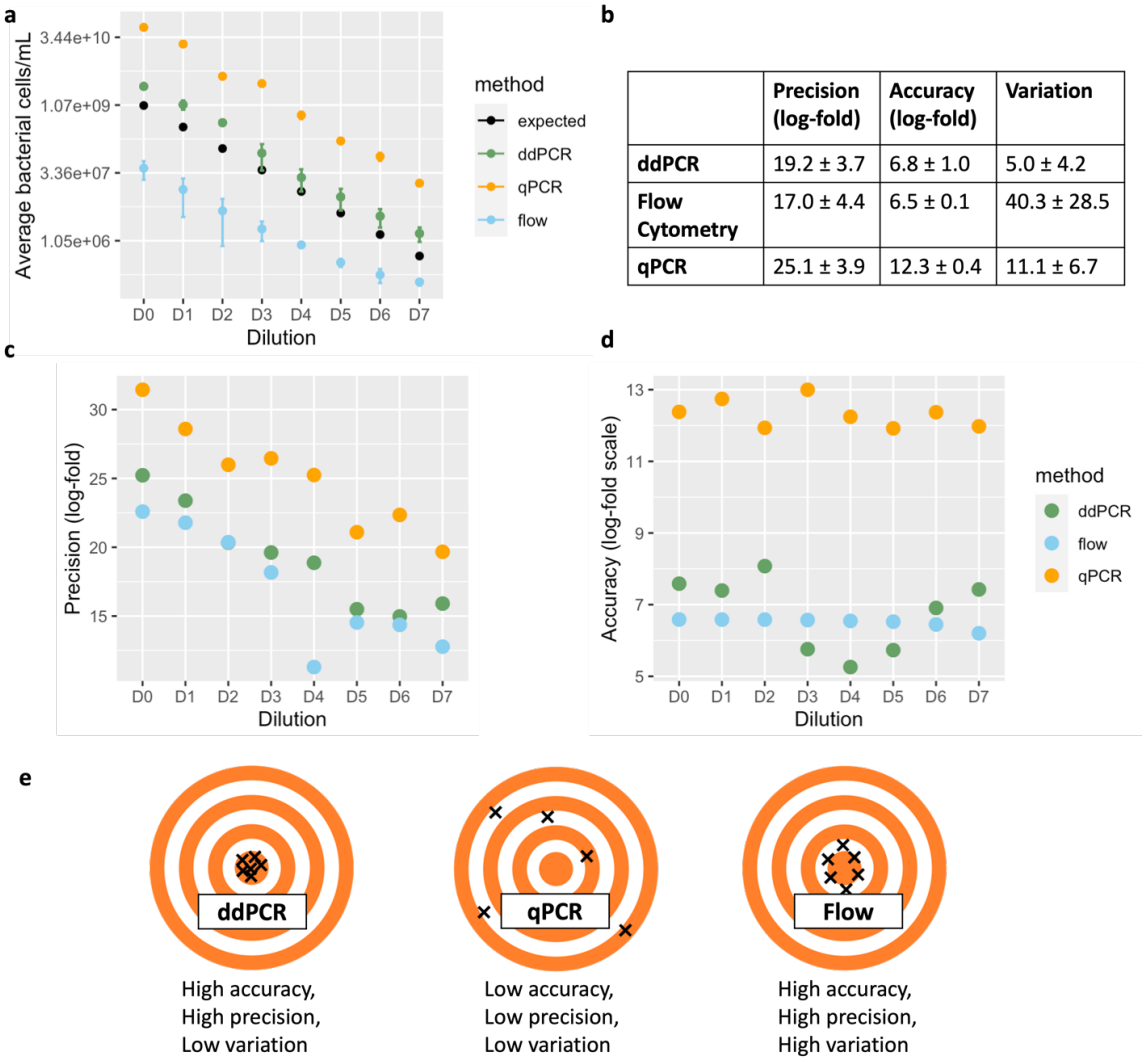
Genera	Genomic DNA (%)	Cell Number (%)	16S genes (%)
Pseudomonas	12	6.1	4.2
Escherichia	12	8.5	10.1
Salmonella	12	8.8	10.4
Lactobacillus	12	21.9	18.4
Enterococcus	12	14.6	9.9
Staphylococcus	12	15.3	15.5
Listeria	12	13.9	14.1
Bacillus	12	10.3	17.4

Figure 4.1. Overview of the mock microbial community dilution series.
a) Adapted from Karstens, *et al* (2019), a schematic showing the dilution series and naming of each sample. **b**) A table describing the theoretical composition of the mock community sample. The eight genera should be equally present by genomic DNA, however varies by cell number and 16S gene copies.

Quantitative methods

We first performed the quantitative methods of qPCR, ddPCR, and flow cytometry, then evaluated for precision and accuracy. The mock microbial dilution samples for ddPCR and qPCR samples were run in triplicate at 1:100 and 1:1000 dilutions due to the large variance in copy number by dilution and the machines detection ranges,

while the samples run on flow cytometry were in triplicate. **Figure 4.2a** displays an overview of the quantitative results compared to the expected bacterial cells per mL. We found that ddPCR and qPCR had low variation in their replicates but flow had high variation (**Figure 4.2a-b**). Our results demonstrate that both ddPCR and flow had high accuracy and high precision (**Figure 4.2c-d**), however qPCR struggled with both. Noticeably, flow had no change in accuracy regardless of dilution (**Figure 4.2d**), however ddPCR had the best results with the middle dilutions, and worse accuracy on either end of the dilution series.



	Precision (log-fold)	Accuracy (log-fold)	Variation
ddPCR	19.2 ± 3.7	6.8 ± 1.0	5.0 ± 4.2
Flow Cytometry	17.0 ± 4.4	6.5 ± 0.1	40.3 ± 28.5
qPCR	25.1 ± 3.9	12.3 ± 0.4	11.1 ± 6.7

Figure 4.2. Overview of ddPCR (green), qPCR (orange), and flow (blue) measurements compared to expected (black) for average bacterial cells/mL at each dilution.

a) Scatter plot demonstrating the variance in methods compared to expected. **b**) Summary of the log-fold of precision, log-fold change in precision, and variation of each method. **c**) Scatter plots of the precision and **d**) accuracy of each method by dilution. **e**) Schematic summarizing how well the quantitative methods did in terms of accuracy, precision, and variation.

Sequencing overview

We next evaluated the composition of the mock microbial dilution series with 16S sequencing before combining with the quantitative methods to perform QMP. The eight dilutions from the mock microbial dilution series were sequenced and eight known genera identified in each sample (**Figure 4.3a-b**). The average reads of the mock community dilutions were 166578.9 per sample (interquartile range 152127.5, 187004.8 reads). Interestingly, D4 has the most reads while D7 has the least (**Figure 4.3b**). Despite having only eight expected genera from two phyla, we identified 11 phyla and 73 genera. The inflated phyla and genera is most likely from contaminating sequences, which is typically seen as microbial biomass decreases.[65]

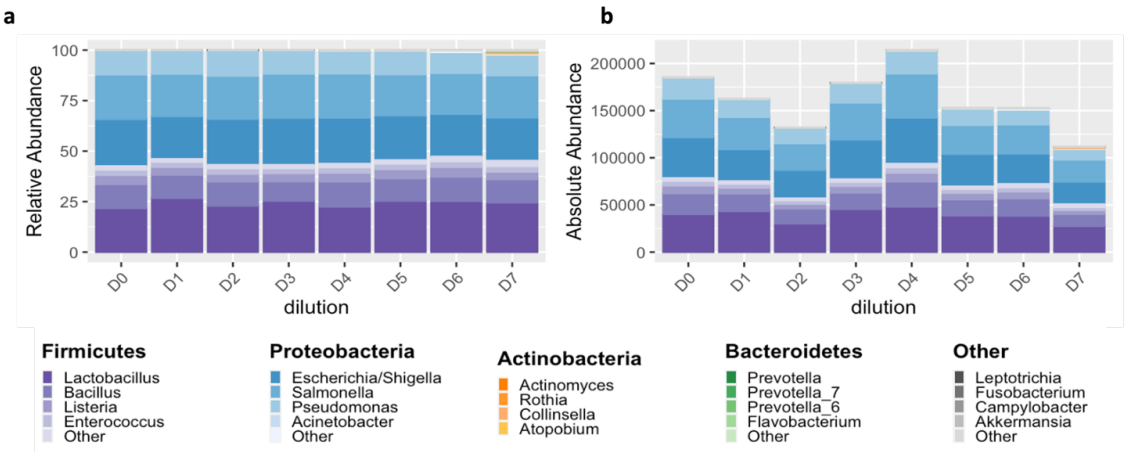


Figure 4.3. Sequencing overview of the mock microbial dilution series.
a) Relative abundance and **b**) raw read count stacked bar plots of the mock microbial dilution series. The expected genera are Firmicutes (purple) and Proteobacteria (blue).

Applying QMP

To apply QMP, we combined quantitative measures of microbial load (ddPCR, flow, and qPCR) with the sequencing data. This involves rarefying the samples based on a ratio of the sample's sequencing depth to its cell counts, and then multiplying by the cell counts.[155] We then compared the resulting QMP abundance from each method to the expected cell counts (**Figure 4.4**). The composition of the qPCR samples at D0 through D4 are distinct from the expected composition as well as the other methods. Overall, both flow and ddPCR looked similar to the expected at all concentrations of the mock community dilution series, however ddPCR had less *Pseudomonas* in D0 than flow and the expected.

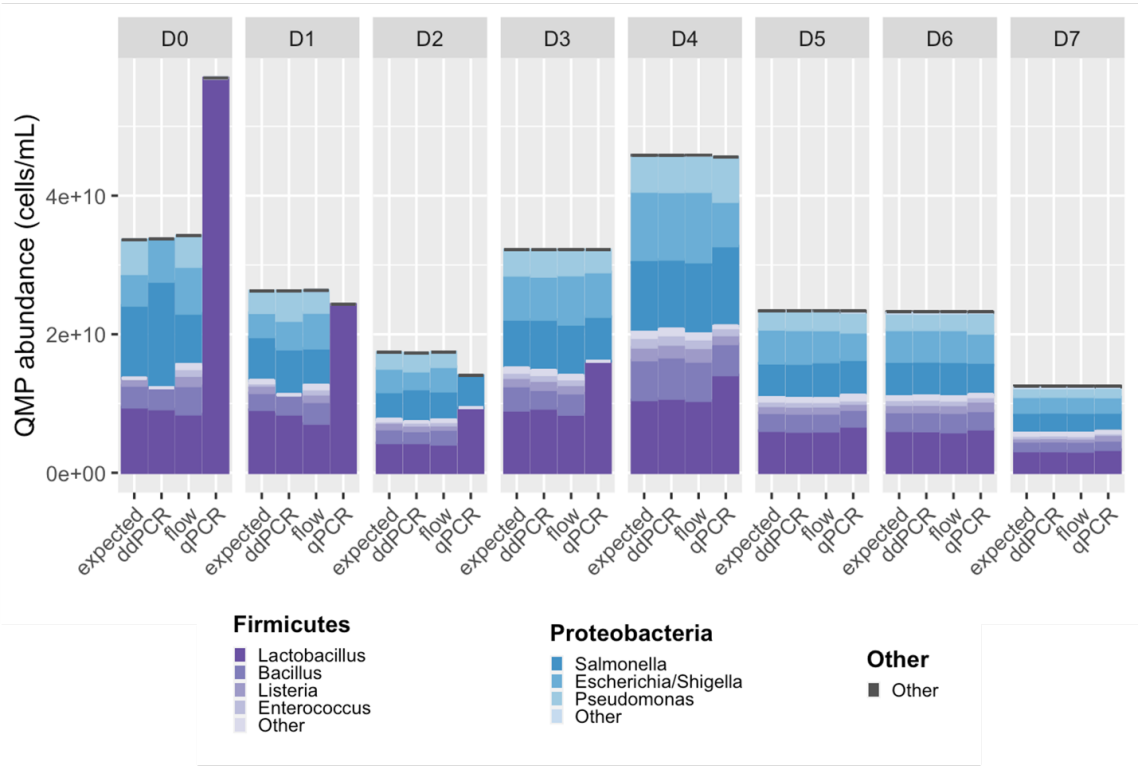


Figure 4.4 Stacked bar plots of each quantitative method for each dilution from the mock microbial dilution series using QMP abundance.

4.4 Discussion

Our results demonstrated that ddPCR and flow performed better than qPCR with regards to accuracy and precision in measuring bacterial cells in a sample (**Figure 2b**). Although flow cytometry did well with precision and accuracy, it is generally more expensive and takes certain expertise as compared to both ddPCR and qPCR. Many more labs have ddPCR and qPCR machines and running multiple replicates is feasible, whereas flow cytometers are often a shared resource between many labs.

We found that applying the QMP method was straightforward, however there were some issues with qPCR. As the qPCR bacterial cell counts for D0 were so high (**Figure 4.2a**), rarefying by the ratio of sequencing depth to cell counts meant we had to rarefy D0 to 0.60 sequencing reads. This doesn't make biological sense, therefore we rarefied to one read, which also has issues. We believe this means researchers should be careful about using this specific QMP method and qPCR when studying high-microbial biomass samples.

Many different QMP methods have been published since this work began. Some QMP methods rely on knowing what will be in your sample – something which is not always possible when surveying a less characterized microbiome, like the urinary microbiome.[156,161,162] Ideally, QMP methods would be able to be applied to any sample set and not require that prior knowledge. In the future, we would like to benchmark these newer methods on the mock microbial dilution series, as well as apply them to biological samples. Specifically, future work includes applying these methods to high- and low-microbial biomass biological samples, such as fecal and urine samples.

4.5 Methods

DNA Extraction

A commercially available mock microbial community (Zymo, Cat# D6300) was serially diluted per previous research prior to DNA extraction.[65] The mock community has 8 bacterial species (*Pseudomonas aeruginosa*, *Escherichia coli*, *Salmonella enterica*, *Lactobacillus fermentum*, *Enterococcus faecalis*, *Staphylococcus aureus*, *Listeria monocytogenes*, and *Bacillus subtilis*) and was diluted with microbial DNA-free water

(Qiagen) in 8 rounds (D0-D7) of a serial 3-fold dilution. Microbial-free water (Qiagen) was used as negative controls and processed alongside mock community samples. Microbial DNA was extracted using the QiAMP DNA mini kit (Qiagen).

Library Preparation and Sequencing

The extracted DNA was quantified using a NanoDrop spectrophotometer (Thermo Fisher Scientific, USA) prior to PCR amplification. Bacterial DNA was amplified by PCR with Golay primers which target the V4 region of 16S rRNA genes (Forward: GTGCCAGCMGCCGCGGTAA; Reverse: GGACTACHVGGGTWTCTAAT).[64] Template DNA was amplified in triplicate using the GoTaq Hot Start polymerase kit (Promega, USA). Amplified DNA from the mock community and negative controls were submitted for 16S rRNA gene sequencing on an Illumina MiSeq at Oregon State University.

Droplet-Digital PCR Protocol

Each PCR reaction was prepared in a PCR hood in a dedicated room for ddPCR. Each dilution of the mock community was diluted either 1:1000 or 1:100 in microbial-free water (Qiagen). In total, the ddPCR reaction mixture used was 22 µL containing primers, template, and QX200 ddPCR EvaGreen Supermix (Bio-Rad). The QX200 Digital PCR System with Auto Droplet Generator and Reader (Biorad) was used. The PCR protocol uses a 2C/second ramp rate and starts with a 10-minute 95C enzyme activation step followed by 40 cycles of a two-step protocol (94C for 30 seconds and 59C for 1 minute), and lastly cycles up to 98C for 10 minutes. V6

primers from the Sfanos lab at Hopkins were used (Forward: CAACGCGWRGAACCTTACC; Reverse: CRRCACGAGCTGACGAC).

Quantitative PCR Protocol

Previously extracted DNA from the mock community dilution series were used along with PerfeCTa SYBR Green SuperMix (QuantaBio) and measured in triplicate. Forward (ACTCCTACGGGAGGCAGCAGT) and reverse (ATTACCGCGGCTGCTGGC) primers targeting the 16S rRNA gene were used.[163] Triplicates of each sample were run on the ViiA 7 (Applied Biosystems) and resulting data was generated and analyzed using the QuantStudio(TM) Real-Time PCR Software. Denaturing, annealing, extension temperatures and time optimized for the amplification process is shown in **Table 4.1**.

Table 4.1. A table describing the qPCR amplification settings.

	Temperature (°C)	**Time (minutes)**	
Ramp	95	5m	
Denature	95	45s	40 cycles
Annealing	60	30s	
Extension	72	45s	
Final	72	6m	
Rest	8	Infinity	

Flow Cytometry

Flow cytometry was used to count whole bacterial cells from mock microbial dilution samples prior to DNA extraction. These samples were stained with SYTO BC and a known concentration of beads was added to the sample. The mixture is then applied to the flow cytometer (Cytoflex). Each sample was run for 30 seconds and a threshold of 1000 events was applied. Fluorescence events were monitored using the FL1 488-1/640-1 nm optical detector. To exclude background events and obtain an accurate microbial cell count, we evaluated the fluorescence events using the forward–sideways density plot. Instrument and gating settings were identical for all samples.

Bioinformatics

Precision, accuracy, and variation were calculated with the equations in **Figure 4.5**. All analysis was performed in R with the sequencing data agglomerated at the genus level.

Accuracy:

$$\left(\frac{average - expected}{expected}\right) * 100$$

Precision:

$$\frac{\sqrt{\frac{\sum(x - \bar{x})^2}{n}}}{number\ of\ measurements}$$

Variation:

$$\left(\frac{\sqrt{\frac{\sum(x - \bar{x})^2}{n}}}{average}\right) * 100$$

Figure 4.5. Equations for accuracy, precision, and variation.

Chapter 5. Discussion

5.1 Summary and Key Findings

Prostate disease is ubiquitous in aging men.[5,164] Prostate disease includes BPH, PC, and prostatitis, and these diseases share many symptoms such as LUTS and blood in urine.[4,6,18] There have been efforts to develop better prostate disease tests, including interrogating the composition of bacteria in urine and in blood. In the three studies above, we evaluated the urinary microbiome and its relationships with LUTS and prostate disease in a large cohort of men, evaluated microbial DNA in fractions of blood of men with PC, and benchmarked microbiome analysis methods.

The studies in which we examined microbes in urine and blood in the context of prostate disease have yielded several main findings. Starting with the urinary microbiome (**Chapter 2**) the findings are as follows: 1) The composition of the urinary microbiome is associated with BPH, BMI, and age, 2) Specifically, *Dialister* is more prevalent in the urinary microbiomes of obese men than overweight and healthy men, and 3) LUTS symptoms are not only dependent on BMI and the diversity of the urinary microbiome, but also whether the individual has BPH.

In our study of circulating microbial DNA in men with prostate cancer (**Chapter 3**), we discovered: 1) The presence of cell-free circulating microbial DNA, 2.) RBCs have the highest microbial diversity of the blood fractions, 3) There are prostate-disease associated microbial signatures in blood, and 4) There is overlap between the bacteria found in urine and in blood.

We believe the urinary microbiome and microbial DNA in blood both have the potential to be informative with respect to *some* prostate diseases. In our urinary microbiome study, 22.1% of men had a history of prostatitis, and rates of prostatitis were higher in the men with more severe LUTS (**Table 2.1**). Our study, however, was not able to detect any association with the urinary microbiome. For BPH, almost half of the participants had BPH with significantly more men in the more severe LUTS having a BPH diagnosis. As mentioned earlier, we determined the urinary microbiome diversity to be associated with BPH and identified a complex association with LUTS and BMI. Only 13.7% of men in the urinary microbiome study had a history of PC, which could explain why no associations were identified. We did however establish circulating microbial DNA compositional differences in men with and without PC. So while associations were discovered between the human microbiome and both BPH and PC, more work is needed to determine if the same is true for prostatitis.

Interpretation of Results

The DNase-treated samples were invaluable for getting closer to addressing where microbial DNA is in blood (**Figure 3.2b**), and we did not expect the majority of the microbial DNA to be cell-free. One of the many ways the innate immune system responds to infectious agent is by sensing microbial DNA and degrading it.[165] Due to this, we had believed that the immune system would have "cleaned up" the extracellular DNA. Instead, microbial DNA being cell-free could indicate that these bacteria recently entered circulation and the immune system had not yet activated or had a chance to degrade the microbial DNA. Another potential explanation is there is immune dysregulation in the men with disease, however we did not detect

differences in copies of 16S by disease status. Although we were able to measure copies of 16S for all DNase-treated samples, we did not submit all samples for sequencing and DNase samples often failed to generate reads. This limited the conclusions we could make regarding the composition of these samples.

It's clear from our results that BMI has a strong association and potentially impacts the urinary microbiome composition in men. Specifically, we demonstrated a complex relationship between BMI, BPH, irritative and obstructive LUTS, and alpha diversity of the urinary microbiome (**Figure 2.3e-f**). Our results show that within the healthy weight and overweight populations, presence of BPH impacts the trend direction for the irritative and obstructive symptoms with respect to alpha diversity. However, BPH no longer has that effect within the obese population. This could be because at higher BMIs, the weight has a larger impact on the composition of the urinary microbiome than the effect of the enlarged prostate from BPH. Unfortunately, very few studies have examined the relationship of BMI with the male urinary microbiome composition, however BMI has been documented to be associated with changes in the gut and oral microbiomes.[115,116,166]

With respect to *Dialister*, we postulate that there may be metabolites in the bladders of obese men that *Dialister* utilizes, like succinate. *Dialister* is known as a "succinate-consumer" and it uses succinate as nutrients as opposed to carbohydrates.[167,168] Succinate is a metabolite which is associated with impaired metabolism and has been found elevated in the circulation of obese individuals.[169] Succinate has also been found in urine under conditions of stress and inflammation.[170] Interestingly, one study found the bladder to be a primary target of succinate due to the amount of succinate receptors in the urothelial cells.[171] Another study using rats discovered that in rats with hypertension, succinate decreased bladder function and led to LUTS, however the mechanism is unclear.[172]

We believe succinate should be investigated further as potentially playing a role in altering the urinary microbiome composition especially in men with metabolic disorders.

We did identify *Dialister* in the blood and matched urine sample of one patient who was diagnosed with high-grade PC (**Chapter 3**). Interestingly, *Dialister* was found in a DNase-treated buffy coat sample meaning it is most likely intracellular DNA, however *Dialister* was not present in the non-DNase sample (**Figure 3.5c**, patient 4). This patient had a BMI of 25.3 at time of sample collection, which is technically overweight, however we have no other clinical information regarding the possibility of a metabolic disorder.

We were not surprised to find there was overlap between the bacteria found in the blood and the bacteria found in urine of matched patients, because the prostate shares half of its blood supply with the bladder.[173] We sequenced matched urine samples for 14 out of 15 men and found the most common shared genera were *Lactobacillus*, *Prevotella*, *Bacillus*, and *Staphylococcus*. Additionally, we found that 11 of the 15 shared genera between blood and urine is from the Firmicutes phyla, and Firmicutes was the most abundant overlapping phyla regardless of disease status. Cavaretta et al. found Firmicutes to be the second most abundant phylum in prostate tissue,[174] so one explanation could be that the bacteria are coming from the bladder and colonizing the prostate.

5.2 Ongoing Technical Considerations

Many of the most common microbiome analysis methods have been optimized for high microbial biomass samples, such as stool. As discussed in **Chapter 4**, often what

works well for high microbial biomass samples is not appropriate for low microbial biomass samples. Overall, there is no clear guidance on how best to handle low microbial biomass samples and associated contamination. There has been effort into publishing the most common contaminants introduced by the "kitome",[175,176] and new computational methods of identifying contaminants,[177,178] however contamination is seemingly personalized to each laboratory environment and changes over time.[175,179]

Despite our many positive and negative controls, we experienced challenges with contamination in our work. Namely, a portion of the reads generated from sequencing our blood samples were from marine bacteria and these bacteria were not always present in our controls. Anecdotally, we have however discovered that using full-length 16S sequencing on our low microbial biomass samples decreases the overall contaminants in our samples and often our negative controls will fail to generate reads. Part of the pre-processing of full-length sequencing data is to remove any sequences below 1500 bps – or the length of the 16S gene – therefore we believe the majority of contaminant DNA is fragmented. We believe using full-length 16S sequencing could be the way forward for low microbial biomass samples, however full-length may only be appropriate for sample types in which bacteria is intact and will thus have long DNA sequences for successful 16S full-length sequencing.

Although the benefit of using full-length 16S sequencing is handling of contamination, we had quality issues with our blood fraction samples (**Chapter 3**). We noticed significant amounts of short reads (<400 bps) generated from the full-length 16S sequencing of blood fractions, and very few short reads from our mock microbial dilution series. The DNA from the mock microbial dilution series is isolated from intact microbial cells which could explain the discrepancy with

sequencing read length. We have found that the short DNA sequences do correspond to known microbes. We hypothesize the short sequencing reads in the blood fractions indicate the microbial DNA in blood is fragmented and there may not be intact microbes. More work is needed to investigate the short sequencing reads and if they are from human-associated microbes.

Additional issues in the microbiome field to consider is the reliance on taxonomic databases, which can cause many different problems. Taxonomic databases are often updated every few years, which while helpful, can cause issues because taxonomic names are changing constantly. For example, I used an updated version of the SILVA database (138.1) and was shocked when I could not find *Lactobacillus* in my positive controls, when in fact, *Lactobacillus* had been renamed *Lactobacillaceae* in this newer version of the database. Not only does this cause an issue for reproducibility, but also for researching a specific microbe by name across many years of studies.

5.3 Next Steps and Future Directions

First, we will discuss specific next steps for our projects using urine and blood to better understand aspects of prostate disease. Then we will mention exciting future directions for the field.

Specific Next Steps

In our studies, we examined the urinary microbiome and microbial DNA in blood in the context of prostate disease. Although associations were found with BPH and

PC, there was not enough men in our urinary microbiome study cohort who had prostatitis to draw conclusions on associations with the urinary microbiome (**Chapter 2**). It would be beneficial to have an even larger cohort for evaluating associations with prostatitis and/or PC, especially given the heterogeneity of the urinary microbiome. Fortunately, there are at least one thousand more urine samples from the men of the MrOS cohort (**Chapter 2**) which have been collected at many time points and there is data available on if these men went on to eventually develop PC. This cohort of men has been heavily genotyped and phenotyped, which includes medication and diet data – neither of which have been explored to have an impact on the urinary microbiome. Prior work has demonstrated diet impacts risk of developing urinary tract infections, therefore diet is most likely an important covariate to consider.[180] We believe using full-length 16S sequencing on these samples would allow us to tackle a few goals: 1) Determine urinary microbiome stability over time – important for putting previous work in context, 2) test associations between urinary microbiome composition and prostatitis, 3) Begin to unravel the complex relationships between BPH, BMI, LUTS, and urinary microbiome diversity.

The next steps for the interrogation of circulating microbial DNA (**Chapter 3**) would be to first explore the short sequencing reads mentioned previously. This would include evaluating each of the controls using only long read (>1500 bps) versus only short reads (<1500) and observing how the microbial composition changes. We would repeat this for all samples for all fractions. It will also involve a literature review of the microbes assigned to the short sequencing reads and if they are known to be human associated. There is potential that these short sequencing reads will change our interpretation of the data, so pending those results, we would scale up our blood fraction study. Specifically, more investigation is needed into the

RBC fraction and how it has the highest alpha diversity of the other fractions, and how this could provide the most information about microbial DNA in blood.

Future Directions

The landscape for human microbiome research in prostate disease is vast. We are particularly excited about the work determining the impact of other microbial members in circulation and the urogenital tract, such as fungus and phage.[181,182] The mycobiome has yet to be charactered in the bladder and efforts to characterize phage are underway.[183] Another exciting direction for microbiome research is exploring the role hormones play in both affecting microbial communities and disease. For instance, higher levels of testosterone, which is known to drive PC,[184] have been associated with abundance of Firmicutes in the gut microbiome,[185] and microbes in the gut have been found to produce androgens and facilitate the development of treatment resistant-PC.[186] There have already been studies of the excreted androgens through the urine with respect to PC,[187] however no research into the effect on the urinary microbiome.

Thinking more broadly, we believe we can also look to the most recent gut microbiome research for new ways to interrogate low microbial biomass samples. Studies on the gut microbiome have been moving towards functional studies and microbiome genetics. In contrast, many of the studies on the urinary microbiome and circulating microbial DNA have been association-based, and better understanding of the functions of microbial communities is needed. We believe moving towards metagenomic sequencing and combining that data with metabolomics holds great promise to move the low microbial biomass field forward.

In conclusion, these lesser studied microbial niches in the human body are a rich area for exploration in their interactions with prostate health. We feel the studies presented have identified key elements for consideration in study design, sample processing, and data handling as well as generated new areas of focus. We look forward to continued exploration in this field and to see how others build on this body of knowledge.

References

1. Ford, E. Case of a Diseased Bladder, with Enlargement of the Prostate Gland, and Other Morbid Appearances. *Med Phys J* **4**, 390–393 (1800).
2. Crowther, C. Case of Induration of the Vesiculæ Seminales, Prostate Glands, and Rectum. *Med Phys J* **4**, 506–508 (1800).
3. Berry, S. J., Coffey, D. S., Walsh, P. C. & Ewing, L. L. The Development of Human Benign Prostatic Hyperplasia with Age. *Journal of Urology* **132**, 474–479 (1984).
4. Sarma, A. V. & Wei, J. T. Benign Prostatic Hyperplasia and Lower Urinary Tract Symptoms. *New England Journal of Medicine* **367**, 248–257 (2012).
5. Siegel, R. L., Miller, K. D., Wagle, N. S. & Jemal, A. Cancer statistics, 2023. *CA Cancer J Clin* **73**, 17–48 (2023).
6. Sandhu, S. *et al.* Prostate cancer. *The Lancet* **398**, 1075–1090 (2021).
7. Nickel, J. C., Downey, J., Hunter, D. & Clark, J. Prevalence of Prostatitis-like Symptoms in a Population Based Study Using the National Institutes of Health Chronic Prostatitis Symptom Index. *Journal of Urology* **165**, 842–845 (2001).
8. Wagenlehner, F. *et al.* Prostatitis and andrological implications. *Minerva Urol Nefrol* **65**, 117–23 (2013).
9. Krieger, J. N., Nyberg Leroy, J. & Nickel, J. C. NIH Consensus Definition and Classification of Prostatitis. *JAMA* **282**, 236–237 (1999).
10. Madersbacher, S., Marszalek, M., Lackner, J., Berger, P. & Schatzl, G. The Long-Term Outcome of Medical Therapy for BPH. *Eur Urol* **51**, 1522–1533 (2007).
11. Austin, D. C. *et al.* NF-κB and androgen receptor variant expression correlate with human BPH progression. *Prostate* **76**, 491–511 (2016).

12. Traish, A. M. The Impact of the 5α-Reductase Inhibitors (5α-RIs) on Male Sexual Function and Psychological Well-Being. *Curr Sex Health Rep* **7**, 210–219 (2015).

13. Ottaiano, N., Shelton, T., Sanekommu, G. & Benson, C. R. Surgical Complications in the Management of Benign Prostatic Hyperplasia Treatment. *Curr Urol Rep* **23**, 83–92 (2022).

14. Litwin, M. S. & Tan, H.-J. The Diagnosis and Treatment of Prostate Cancer. *JAMA* **317**, 2532 (2017).

15. Khan, F. U. *et al.* Comprehensive overview of prostatitis. *Biomedicine & Pharmacotherapy* **94**, 1064–1076 (2017).

16. Kably, I., Acharya, V., Richardson, A. J. & Bhatia, S. Prostatic Artery Embolization in Refractory Hematuria of Prostatic Origin. *Tech Vasc Interv Radiol* **23**, 100694 (2020).

17. Hamilton, W., Sharp, D. J., Peters, T. J. & Round, A. P. Clinical features of prostate cancer before diagnosis: a population-based, case-control study. *Br J Gen Pract* **56**, 756–62 (2006).

18. Schaeffer, A. J. Chronic Prostatitis and the Chronic Pelvic Pain Syndrome. *New England Journal of Medicine* **355**, 1690–1698 (2006).

19. Laborde, E. E. & McVary, K. T. Medical management of lower urinary tract symptoms. *Rev Urol* **11**, S19-25 (2009).

20. Ganz, M. L. *et al.* Economic Costs of Overactive Bladder in the United States. *Urology* **75**, 526-532.e18 (2010).

21. WEI, J. T., CALHOUN, E. & JACOBSEN, S. J. Urologic Diseases in America Project: Benign Prostatic Hyperplasia. *Journal of Urology* **173**, 1256–1261 (2005).

22. PDQ Screening and Prevention Editorial Board. *Prostate Cancer Prevention (PDQ®): Patient Version*. (2002).

23. Kuriyama, M. *et al.* Quantitation of Prostate-specific Antigen in Serum by a Sensitive Enzyme Immunoassay1. *Cancer Res* **40**, 4658–4662 (1980).

24. Merriel, S. W. D., Funston, G. & Hamilton, W. Prostate Cancer in Primary Care. *Adv Ther* **35**, 1285–1294 (2018).

25. Ilic, D., Neuberger, M. M., Djulbegovic, M. & Dahm, P. Screening for prostate cancer. *Cochrane Database Syst Rev* **2013**, CD004720 (2013).

26. Bohnen, A. M., Groeneveld, F. P. & Bosch, J. L. H. R. Serum Prostate-Specific Antigen as a Predictor of Prostate Volume in the Community: The Krimpen Study. *Eur Urol* **51**, 1645–1653 (2007).

27. Gamé, X. *et al.* Total and Free Serum Prostate Specific Antigen Levels during the First Month of Acute Prostatitis. *Eur Urol* **43**, 702–705 (2003).

28. Pansadoro, V. *et al.* Prostate-Specific Antigen and Prostatitis in Men under Fifty. *Eur Urol* **30**, 24–27 (1996).

29. Lepor, H. Evaluating men with benign prostatic hyperplasia. *Rev Urol* **6 Suppl 1**, S8–S15 (2004).

30. Soekmadji, C. *et al.* Extracellular vesicles for personalized therapy decision support in advanced metastatic cancers and its potential impact for prostate cancer. *Prostate* **77**, 1416–1423 (2017).

31. Øverbye, A. *et al.* Identification of prostate cancer biomarkers in urinary exosomes. *Oncotarget* **6**, 30357–30376 (2015).

32. Tavoosidana, G. *et al.* Multiple recognition assay reveals prostasomes as promising plasma biomarkers for prostate cancer. *Proceedings of the National Academy of Sciences* **108**, 8809–8814 (2011).

33. Boddy, J. L., Gal, S., Malone, P. R., Harris, A. L. & Wainscoat, J. S. Prospective Study of Quantitation of Plasma DNA Levels in the Diagnosis of Malignant versus Benign Prostate Disease. *Clinical Cancer Research* **11**, 1394–1399 (2005).

34. Gordian, E. *et al.* Serum Free Circulating DNA Is a Useful Biomarker to Distinguish Benign versus Malignant Prostate Disease. *Cancer Epidemiology, Biomarkers & Prevention* **19**, 1984–1991 (2010).

35. Wagenlehner, F. M. E. *et al.* The role of inflammation and infection in the pathogenesis of prostate carcinoma. *BJU Int* **100**, 733–737 (2007).

36. Ficarra, V. *et al.* The Role of Inflammation in Lower Urinary Tract Symptoms (LUTS) due to Benign Prostatic Hyperplasia (BPH) and Its Potential Impact on Medical Therapy. *Curr Urol Rep* **15**, 463 (2014).

37. Sfanos, K. S., Yegnasubramanian, S., Nelson, W. G. & De Marzo, A. M. The inflammatory microenvironment and microbiome in prostate cancer development. *Nat Rev Urol* **15**, 11–24 (2018).

38. COHEN, R. J., SHANNON, B. A., McNEAL, J. E., SHANNON, T. & GARRETT, K. L. Propionibacterium acnes Associated with Inflammation in Radical Prostatectomy Specimens: A Possible Link to Cancer. *Journal of Urology* **173**, 1969–1974 (2005).

39. Sender, R., Fuchs, S. & Milo, R. Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. *Cell* **164**, 337–340 (2016).

40. Gilbert, J. A. *et al.* Current understanding of the human microbiome. *Nat Med* **24**, 392–400 (2018).

41. Rinninella, E. *et al.* What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. *Microorganisms* **7**, 14 (2019).

42. Gill, S. R. *et al.* Metagenomic Analysis of the Human Distal Gut Microbiome. *Science (1979)* **312**, 1355–1359 (2006).

43. Lee, Y. K. & Mazmanian, S. K. Has the Microbiota Played a Critical Role in the Evolution of the Adaptive Immune System? *Science (1979)* **330**, 1768–1773 (2010).

44. Sheh, A. & Fox, J. G. The role of the gastrointestinal microbiome in Helicobacter pylori pathogenesis. *Gut Microbes* **4**, 505–531 (2013).

45. Sanchez-Rodriguez, E. *et al.* The Gut Microbiota and Its Implication in the Development of Atherosclerosis and Related Cardiovascular Diseases. *Nutrients* **12**, 605 (2020).

46. de Groot, P. F. *et al.* Distinct fecal and oral microbiota composition in human type 1 diabetes, an observational study. *PLoS One* **12**, e0188475 (2017).

47. Hilty, M. *et al.* Disordered Microbial Communities in Asthmatic Airways. *PLoS One* **5**, e8578 (2010).

48. Jiang, H. *et al.* Altered fecal microbiota composition in patients with major depressive disorder. *Brain Behav Immun* **48**, 186–194 (2015).

49. Schwabe, R. F. & Jobin, C. The microbiome and cancer. *Nat Rev Cancer* **13**, 800–812 (2013).

50. Golombos, D. M. *et al.* The Role of Gut Microbiome in the Pathogenesis of Prostate Cancer: A Prospective, Pilot Study. *Urology* **111**, 122–128 (2018).

51. Dohnalová, L. *et al.* A microbiome-dependent gut–brain pathway regulates motivation for exercise. *Nature* **612**, 739–747 (2022).

52. Ning, L. *et al.* Microbiome and metabolome features in inflammatory bowel disease via multi-omics integration analyses across cohorts. *Nat Commun* **14**, 7135 (2023).

53. Chen, X., Lu, Y., Chen, T. & Li, R. The Female Vaginal Microbiome in Health and Bacterial Vaginosis. *Front Cell Infect Microbiol* **11**, (2021).

54. Peng, X. *et al.* Oral microbiota in human systematic diseases. *Int J Oral Sci* **14**, 14 (2022).

55. Byrd, A. L., Belkaid, Y. & Segre, J. A. The human skin microbiome. *Nat Rev Microbiol* **16**, 143–155 (2018).

56. Torres, P. J. *et al.* Characterization of the salivary microbiome in patients with pancreatic cancer. *PeerJ* **3**, e1373 (2015).

57. Urbaniak, C. *et al.* Microbiota of human breast tissue. *Appl Environ Microbiol* **80**, 3007–14 (2014).

58. Banerjee, S. *et al.* Microbiome signatures in prostate cancer. *Carcinogenesis* **40**, 749–764 (2019).

59. Nejman, D. *et al.* The human tumor microbiome is composed of tumor type–specific intracellular bacteria. *Science (1979)* **368**, 973–980 (2020).

60. Blum, H. E. The human microbiome. *Adv Med Sci* **62**, 414–420 (2017).

61. Kustrimovic, N., Bombelli, R., Baci, D. & Mortara, L. Microbiome and Prostate Cancer: A Novel Target for Prevention and Treatment. *Int J Mol Sci* **24**, (2023).

62. Neugent, M. L., Hulyalkar, N. V., Nguyen, V. H., Zimmern, P. E. & De Nisco, N. J. Advances in Understanding the Human Urinary Microbiome and Its Potential Role in Urinary Tract Infection. *mBio* **11**, (2020).

63. Kolbert, C. P. & Persing, D. H. Ribosomal DNA sequencing as a tool for identification of bacterial pathogens. *Curr Opin Microbiol* **2**, 299–305 (1999).

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

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

66. Stämmler, F. *et al.* Adjusting microbiome profiles for differences in microbial load by spike-in bacteria. *Microbiome* **4**, 28 (2016).

67. Willis, A. D. Rarefaction, Alpha Diversity, and Statistics. *Front Microbiol* **10**, (2019).

68. Beta-diversity distance matrices for microbiome sample size and power calculations — How to obtain good estimates. *Comput Struct Biotechnol J* **20**, 2259–2267 (2022).

69. Lewis, D. A. *et al.* The human urinary microbiome; bacterial DNA in voided urine of asymptomatic adults. *Front Cell Infect Microbiol* **3**, (2013).

70. 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* **52**, 871–876 (2014).

71. Siddiqui, H., Nederbragt, A. J., Lagesen, K., Jeansson, S. L. & Jakobsen, K. S. Assessing diversity of the female urine microbiota by high throughput sequencing of 16S rDNA amplicons. *BMC Microbiol* **11**, 244 (2011).

72. Karstens, L. *et al.* Does the Urinary Microbiome Play a Role in Urgency Urinary Incontinence and Its Severity? *Front Cell Infect Microbiol* **6**, (2016).

73. Pearce, M. M. *et al.* The female urinary microbiome: A comparison of women with and without urgency urinary incontinence. *mBio* **5**, (2014).

74. Thomas-White, K. J. *et al.* Incontinence medication response relates to the female urinary microbiota. *Int Urogynecol J* **27**, 723–733 (2016).

75. Yu, H. *et al.* Urinary microbiota in patients with prostate cancer and benign prostatic hyperplasia. *Archives of Medical Science* **2**, 385–394 (2015).

76. Shrestha, E. *et al.* Profiling the Urinary Microbiome in Men with Positive versus Negative Biopsies for Prostate Cancer. *Journal of Urology* **199**, 161–171 (2018).

77. Damgaard, C. *et al.* Viable Bacteria Associated with Red Blood Cells and Plasma in Freshly Drawn Blood Donations. *PLoS One* **10**, e0120826 (2015).

78. Tan, C. C. S. *et al.* No evidence for a common blood microbiome based on a population study of 9,770 healthy humans. *Nat Microbiol* **8**, 973–985 (2023).

79. Païssé, S. *et al.* Comprehensive description of blood microbiome from healthy donors assessed by 16 <scp>S</scp> targeted metagenomic sequencing. *Transfusion (Paris)* **56**, 1138–1147 (2016).

80. Sato, J. *et al.* Gut Dysbiosis and Detection of "Live Gut Bacteria" in Blood of Japanese Patients With Type 2 Diabetes. *Diabetes Care* **37**, 2343–2350 (2014).

81. Serena, G., Davies, C., Cetinbas, M., Sadreyev, R. I. & Fasano, A. Analysis of blood and fecal microbiome profile in patients with celiac disease. *Hum Microb J* **11**, 100049 (2019).

82. Shah, N. B. *et al.* Blood Microbiome Profile in CKD. *Clinical Journal of the American Society of Nephrology* **14**, 692–701 (2019).

83. Poore, G. D. *et al.* Microbiome analyses of blood and tissues suggest cancer diagnostic approach. *Nature* **579**, 567–574 (2020).

84. Traykova, D., Schneider, B., Chojkier, M. & Buck, M. Blood Microbiome Quantity and the Hyperdynamic Circulation in Decompensated Cirrhotic Patients. *PLoS One* **12**, e0169310 (2017).

85. Albiger, B., Dahlberg, S., Henriques-Normark, B. & Normark, S. Role of the innate immune system in host defence against bacterial infections: focus on the Toll-like receptors. *J Intern Med* **261**, 511–528 (2007).

86. Egevad, L. *et al.* Benign mimics of prostate cancer. *Pathology* **53**, 26–35 (2021).

87. Boddy, J. L., Gal, S., Malone, P. R., Harris, A. L. & Wainscoat, J. S. Prospective Study of Quantitation of Plasma DNA Levels in the Diagnosis of Malignant versus Benign Prostate Disease. *Clinical Cancer Research* **11**, 1394–1399 (2005).

88. Yu, S. H. & Jung, S. Il. The Potential Role of Urinary Microbiome in Benign Prostate Hyperplasia/Lower Urinary Tract Symptoms. *Diagnostics* **12**, 1862 (2022).

89. Bajic, P. *et al.* Male Bladder Microbiome Relates to Lower Urinary Tract Symptoms. *Eur Urol Focus* **6**, 376–382 (2020).

90. Ahn, H. K., Kim, K., Park, J. & Kim, K. H. Urinary microbiome profile in men with genitourinary malignancies. *Investig Clin Urol* **63**, 569 (2022).

91. Potgieter, M., Bester, J., Kell, D. B. & Pretorius, E. The dormant blood microbiome in chronic, inflammatory diseases. *FEMS Microbiol Rev* **39**, 567–591 (2015).

92. Jing, Y. *et al.* Associations Between Peripheral Blood Microbiome and the Risk of Hypertension. *Am J Hypertens* **34**, 1064–1070 (2021).

93. Sfanos, K. S. *et al.* A molecular analysis of prokaryotic and viral DNA sequences in prostate tissue from patients with prostate cancer indicates the presence of multiple and diverse microorganisms. *Prostate* **68**, 306–320 (2008).

94. Park, S. Y. *et al.* Plasma Microbial Cell-Free DNA Sequencing from over 15,000 Patients Identified a Broad Spectrum of Pathogens. *J Clin Microbiol* (2023) doi:10.1128/jcm.01855-22.

95. Ouaknine Krief, J. *et al.* Role of antibiotic use, plasma citrulline and blood microbiome in advanced non-small cell lung cancer patients treated with nivolumab. *J Immunother Cancer* **7**, 176 (2019).

96. Fouts, D. E. *et al.* Integrated next-generation sequencing of 16S rDNA and metaproteomics differentiate the healthy urine microbiome from asymptomatic bacteriuria in neuropathic bladder associated with spinal cord injury. *J Transl Med* **10**, 174 (2012).

97. Khasriya, R. *et al.* Spectrum of Bacterial Colonization Associated with Urothelial Cells from Patients with Chronic Lower Urinary Tract Symptoms. *J Clin Microbiol* **51**, 2054–2062 (2013).

98. Price, T. *et al.* The urobiome of continent adult women: a cross-sectional study. *BJOG* **127**, 193–201 (2020).

99. Perez-Carrasco, V., Soriano-Lerma, A., Soriano, M., Gutiérrez-Fernández, J. & Garcia-Salcedo, J. A. Urinary Microbiome: Yin and Yang of the Urinary Tract. *Front Cell Infect Microbiol* **11**, (2021).

100. Modena, B. D. *et al.* Changes in Urinary Microbiome Populations Correlate in Kidney Transplants With Interstitial Fibrosis and Tubular Atrophy Documented in Early Surveillance Biopsies. *American Journal of Transplantation* **17**, 712–723 (2017).

101. Gottschick, C. *et al.* The urinary microbiota of men and women and its changes in women during bacterial vaginosis and antibiotic treatment. *Microbiome* **5**, 99 (2017).

102. Vaughan, M. H. *et al.* The Urinary Microbiome in Postmenopausal Women with Recurrent Urinary Tract Infections. *Journal of Urology* **206**, 1222–1231 (2021).

103. Price, T. K. *et al.* Temporal Dynamics of the Adult Female Lower Urinary Tract Microbiota. (2020) doi:10.1128/mBio.

104. Xie, J. *et al.* Profiling the urinary microbiome in men with calcium-based kidney stones. *BMC Microbiol* **20**, 41 (2020).

105. Mansour, B. *et al.* Bladder cancer-related microbiota: examining differences in urine and tissue samples. *Sci Rep* **10**, 11042 (2020).

106. Mondul, A. M., Giovannucci, E. & Platz, E. A. A prospective study of obesity, and the incidence and progression of lower urinary tract symptoms. *Journal of Urology* **191**, 715–721 (2014).

107. Orwoll, E. *et al.* Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study — A large observational study of the determinants of fracture in older men. *Contemp Clin Trials* **26**, 569–585 (2005).

108. Taylor, B. C. *et al.* Prevalence, severity, and health correlates of lower urinary tract symptoms among older men: The MrOS study. *Urology* **68**, 804–809 (2006).

109. De Nunzio, C., Salonia, A., Gacci, M. & Ficarra, V. Inflammation is a target of medical treatment for lower urinary tract symptoms associated with benign prostatic hyperplasia. *World J Urol* **38**, 2771–2779 (2020).

110. Nelson, D. E. *et al.* Characteristic Male Urine Microbiomes Associate with Asymptomatic Sexually Transmitted Infection. *PLoS One* **5**, e14116 (2010).

111. Pinart, M. *et al.* Gut Microbiome Composition in Obese and Non-Obese Persons: A Systematic Review and Meta-Analysis. *Nutrients* **14**, 12 (2021).

112. Ghazi, A. R. *et al.* High-sensitivity pattern discovery in large, paired multiomic datasets. *Bioinformatics* **38**, i378–i385 (2022).

113. Hrbacek, J., Morais, D., Cermak, P., Hanacek, V. & Zachoval, R. Alpha-diversity and microbial community structure of the male urinary microbiota depend on urine sampling method. *Sci Rep* **11**, 23758 (2021).

114. Muñiz Pedrogo, D. A. *et al.* Gut Microbial Carbohydrate Metabolism Hinders Weight Loss in Overweight Adults Undergoing Lifestyle Intervention With a Volumetric Diet. *Mayo Clin Proc* **93**, 1104–1110 (2018).

115. Yan, H. *et al.* Gut Microbiome Alterations in Patients With Visceral Obesity Based on Quantitative Computed Tomography. *Front Cell Infect Microbiol* **11**, (2022).

116. Gao, X. *et al.* Body Mass Index Differences in the Gut Microbiota Are Gender Specific. *Front Microbiol* **9**, (2018).

117. Wang, J. *et al.* Enterotype Bacteroides Is Associated with a High Risk in Patients with Diabetes: A Pilot Study. *J Diabetes Res* **2020**, 1–11 (2020).

118. Holmes, I., Harris, K. & Quince, C. Dirichlet Multinomial Mixtures: Generative Models for Microbial Metagenomics. *PLoS One* **7**, e30126 (2012).

119. Thomas-White, K. J. *et al.* Evaluation of the urinary microbiota of women with uncomplicated stress urinary incontinence. *Am J Obstet Gynecol* **216**, 55.e1-55.e16 (2017).

120. Komesu, Y. M. *et al.* The urinary microbiome in women with mixed urinary incontinence compared to similarly aged controls. *Int Urogynecol J* **29**, 1785–1795 (2018).

121. Li, K. *et al.* Interplay between bladder microbiota and overactive bladder symptom severity: a cross-sectional study. *BMC Urol* **22**, 39 (2022).

122. Burton, E. N. *et al.* Characterization of the urinary microbiome in healthy dogs. *PLoS One* **12**, e0177783 (2017).

123. Colas, L. *et al.* Unique and specific Proteobacteria diversity in urinary microbiota of tolerant kidney transplanted recipients. *American Journal of Transplantation* **20**, 145–158 (2020).

124. Curtiss, N. *et al.* Age, menopausal status and the bladder microbiome. *European Journal of Obstetrics & Gynecology and Reproductive Biology* **228**, 126–129 (2018).

125. Chart, H. Vibrio, mobiluncus, gardnerella and spirillum. in *Medical Microbiology* 314–323 (Elsevier, 2012). doi:10.1016/B978-0-7020-4089-4.00045-7.

126. Pohl, H. G. *et al.* The Urine Microbiome of Healthy Men and Women Differs by Urine Collection Method. *Int Neurourol J* **24**, 41–51 (2020).

127. Liu, R. *et al.* Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. *Nat Med* **23**, 859–868 (2017).

128. Zhang, Q. *et al.* Comparison of gut microbiota between adults with autism spectrum disorder and obese adults. *PeerJ* **9**, e10946 (2021).

129. Cani, P. D. *et al.* Metabolic Endotoxemia Initiates Obesity and Insulin Resistance. *Diabetes* **56**, 1761–1772 (2007).

130. Lee, H. Y. *et al.* The impact of urine microbiota in patients with lower urinary tract symptoms. *Ann Clin Microbiol Antimicrob* **20**, (2021).

131. Li, L.-Y. *et al.* Alterations of gut microbiota diversity, composition and metabonomics in testosterone-induced benign prostatic hyperplasia rats. *Mil Med Res* **9**, 12 (2022).

132. Orwoll, E. *et al.* Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study — A large observational study of the determinants of fracture in older men. *Contemp Clin Trials* **26**, 569–585 (2005).

133. Blank, J. B. *et al.* Overview of recruitment for the osteoporotic fractures in men study (MrOS). *Contemp Clin Trials* **26**, 557–568 (2005).

134. Dahl, E. M., Neer, E., Bowie, K. R., Leung, E. T. & Karstens, L. microshades : An R Package for Improving Color Accessibility and Organization of Microbiome Data. *Microbiol Resour Announc* **11**, (2022).

135. Claassen-Weitz, S. *et al.* Optimizing 16S rRNA gene profile analysis from low biomass nasopharyngeal and induced sputum specimens. *BMC Microbiol* **20**, 113 (2020).

136. Villette, R. *et al.* Refinement of 16S rRNA gene analysis for low biomass biospecimens. *Sci Rep* **11**, 10741 (2021).

137. Callahan, B. J., Grinevich, D., Thakur, S., Balamotis, M. A. & Yehezkel, T. Ben. Ultra-accurate microbial amplicon sequencing with synthetic long reads. *Microbiome* **9**, 130 (2021).

138. Whittle, E., Leonard, M. O., Harrison, R., Gant, T. W. & Tonge, D. P. Multi-Method Characterization of the Human Circulating Microbiome. *Front Microbiol* **9**, 3266 (2018).

139. Yang, D. *et al.* Blood microbiota diversity determines response of advanced colorectal cancer to chemotherapy combined with adoptive T cell immunotherapy. *Oncoimmunology* **10**, (2021).

140. D'Aquila, P. *et al.* Microbiome in Blood Samples From the General Population Recruited in the MARK-AGE Project: A Pilot Study. *Front Microbiol* **12**, (2021).

141. Qiu, J., Zhou, H., Jing, Y. & Dong, C. Association between blood microbiome and type 2 diabetes mellitus: A nested case-control study. *J Clin Lab Anal* **33**, (2019).

142. de Martel, C. *et al.* Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol* **13**, 607–615 (2012).

143. Salter, S. J. *et al.* Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* **12**, 87 (2014).

144. Li, Y. *et al.* Prauserella salsuginis sp. nov., Prauserella flava sp. nov., Prauserella aidingensis sp. nov. and Prauserella sediminis sp. nov., isolated from a salt lake. *Int J Syst Evol Microbiol* **59**, 2923–2928 (2009).

145. Munk, A. C. *et al.* Complete genome sequence of Rhodospirillum rubrum type strain (S1T). *Stand Genomic Sci* **4**, 293–302 (2011).

146. Peña-Castillo, L. *et al.* Gene co-expression network analysis in Rhodobacter capsulatus and application to comparative expression analysis of Rhodobacter sphaeroides. *BMC Genomics* **15**, 730 (2014).

147. Kim, H. *et al.* Bioactive natural products from the genus Salinospora: a review. *Arch Pharm Res* **43**, 1230–1258 (2020).

148. Dekaboruah, E., Suryavanshi, M. V., Chettri, D. & Verma, A. K. Human microbiome: an academic update on human body site specific surveillance and its possible role. *Arch Microbiol* **202**, 2147–2167 (2020).

149. Taylor, S. C., Laperriere, G. & Germain, H. Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: from variable nonsense to publication quality data. *Sci Rep* **7**, 2409 (2017).

150. Minasyan, H. Erythrocyte and blood antibacterial defense. *Eur J Microbiol Immunol (Bp)* **4**, 138–43 (2014).

151. Minasyan, H. Phagocytosis and oxycytosis: two arms of human innate immunity. *Immunol Res* **66**, 271–280 (2018).

152. Corrons, J. L. V., Casafont, L. B. & Frasnedo, E. F. Concise review: how do red blood cells born, live, and die? *Ann Hematol* **100**, 2425–2433 (2021).

153. Abusnina, W., Shehata, M., Karem, E., Koc, Z. & Khalil, E. Clostridium sporogenes bacteremia in an immunocompetent patient. *IDCases* **15**, e00481 (2019).

154. Harith A, A. *et al.* Clostridium Sporogenes Causing Bacteremia Originated from the Skin and Soft Tissue Infection in an Immunocompetent Patient - Case Report and Literature Review. *International Journal of Critical Care and Emergency Medicine* **6**, (2020).

155. Ott, A. *et al.* Improved quantitative microbiome profiling for environmental antibiotic resistance surveillance. *Environ Microbiome* **16**, 21 (2021).

156. McLaren, M. R., Willis, A. D. & Callahan, B. J. Consistent and correctable bias in metagenomic sequencing experiments. *Elife* **8**, (2019).

157. Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V. & Egozcue, J. J. Microbiome Datasets Are Compositional: And This Is Not Optional. *Front Microbiol* **8**, (2017).

158. Pollock, J., Glendinning, L., Wisedchanwet, T. & Watson, M. The Madness of Microbiome: Attempting To Find Consensus "Best Practice" for 16S Microbiome Studies. *Appl Environ Microbiol* **84**, (2018).

159. Kembel, S. W., Wu, M., Eisen, J. A. & Green, J. L. Incorporating 16S Gene Copy Number Information Improves Estimates of Microbial Diversity and Abundance. *PLoS Comput Biol* **8**, e1002743 (2012).

160. Vandeputte, D. *et al.* Quantitative microbiome profiling links gut community variation to microbial load. *Nature* **551**, 507–511 (2017).

161. Jespers, V. *et al.* Quantification of bacterial species of the vaginal microbiome in different groups of women, using nucleic acid amplification tests. *BMC Microbiol* **12**, 83 (2012).

162. Lin, Y., Gifford, S., Ducklow, H., Schofield, O. & Cassar, N. Towards Quantitative Microbiome Community Profiling Using Internal Standards. *Appl Environ Microbiol* **85**, (2019).

163. Scher, J. U. *et al.* Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis. *Elife* **2**, (2013).

164. French, D. B. & Jones, L. A. Minority Issues in Prostate Disease. *Medical Clinics of North America* **89**, 805–816 (2005).

165. Atianand, M. K. & Fitzgerald, K. A. Molecular Basis of DNA Recognition in the Immune System. *The Journal of Immunology* **190**, 1911–1918 (2013).

166. Mohamed Qadir, R. & Assafi, M. S. The association between body mass index and the oral Firmicutes and Bacteroidetes profiles of healthy individuals. *Malays Fam Physician* **16**, 36–43 (2021).

167. Morotomi, M., Nagai, F., Sakon, H. & Tanaka, R. Dialister succinatiphilus sp. nov. and Barnesiella intestinihominis sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol* **58**, 2716–2720 (2008).

168. Sakamoto, M., Ikeyama, N., Iino, T. & Ohkuma, M. Growth of succinate consumer Dialister hominis is supported by Bacteroides thetaiotaomicron. *Anaerobe* **77**, 102642 (2022).

169. Serena, C. *et al.* Elevated circulating levels of succinate in human obesity are linked to specific gut microbiota. *ISME J* **12**, 1642–1657 (2018).

170. Kushnir, M. M., Komaromy-Hiller, G., Shushan, B., Urry, F. M. & Roberts, W. L. Analysis of dicarboxylic acids by tandem mass spectrometry. High-throughput quantitative measurement of methylmalonic acid in serum, plasma, and urine. *Clin Chem* **47**, 1993–2002 (2001).

171. Mossa, A. H., Velasquez Flores, M., Cammisotto, P. G. & Campeau, L. Succinate, increased in metabolic syndrome, activates GPR91 receptor signaling in urothelial cells. *Cell Signal* **37**, 31–39 (2017).

172. Velasquez Flores, M., Mossa, A. H., Cammisotto, P. & Campeau, L. Succinate decreases bladder function in a rat model associated with metabolic syndrome. *Neurourol Urodyn* **37**, 1549–1558 (2018).

173. Singh, O. & Bolla, S. R. *Anatomy, Abdomen and Pelvis, Prostate*. (2023).

174. Cavarretta, I. *et al.* The Microbiome of the Prostate Tumor Microenvironment. *Eur Urol* **72**, 625–631 (2017).

175. Weyrich, L. S. *et al.* Laboratory contamination over time during low-biomass sample analysis. *Mol Ecol Resour* **19**, 982–996 (2019).

176. Glassing, A., Dowd, S. E., Galandiuk, S., Davis, B. & Chiodini, R. J. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. *Gut Pathog* **8**, 24 (2016).

177. Knights, D. *et al.* Bayesian community-wide culture-independent microbial source tracking. *Nat Methods* **8**, 761–763 (2011).

178. Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 226 (2018).

179. Eisenhofer, R. *et al.* Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. *Trends Microbiol* **27**, 105–117 (2019).

180. Chen, Y.-C., Chang, C.-C., Chiu, T. H. T., Lin, M.-N. & Lin, C.-L. The risk of urinary tract infection in vegetarians and non-vegetarians: a prospective study. *Sci Rep* **10**, 906 (2020).

181. Ackerman, A. L. & Underhill, D. M. The mycobiome of the human urinary tract: potential roles for fungi in urology. *Ann Transl Med* **5**, 31–31 (2017).

182. Garretto, A., Miller-Ensminger, T., Wolfe, A. J. & Putonti, C. Bacteriophages of the lower urinary tract. *Nat Rev Urol* **16**, 422–432 (2019).

183. Malki, K. *et al.* Seven Bacteriophages Isolated from the Female Urinary Microbiota. *Genome Announc* **4**, (2016).

184. Michaud, J. E., Billups, K. L. & Partin, A. W. Testosterone and prostate cancer: an evidence-based review of pathogenesis and oncologic risk. *Ther Adv Urol* **7**, 378–387 (2015).

185. Matsushita, M. *et al.* Firmicutes in Gut Microbiota Correlate with Blood Testosterone Levels in Elderly Men. *World J Mens Health* **40**, 517 (2022).

186. Pernigoni, N. *et al.* Commensal bacteria promote endocrine resistance in prostate cancer through androgen biosynthesis. *Science (1979)* **374**, 216–224 (2021).

187. Ventura-Bahena, A. *et al.* Urinary androgens excretion patterns and prostate cancer in Mexican men. *Endocr Relat Cancer* **28**, 745–756 (2021).