

VU Research Portal

Bacterial interactions in the female genital tract

Singer, M.

2019

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Singer, M. (2019). *Bacterial interactions in the female genital tract: A triangle affair between pathogens, microbiota, and host.*

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

CHAPTER 8

The profiling of microbiota in vaginal and urine samples using 16s rRNA gene sequencing and IS-pro analysis

Martin Singer^{1,2}, Rivka Koedooder³, Martine P. Bos⁴, Linda Poort⁵, Paul H.M. Savelkoul^{6,7}, Joop S.E. Laven³, Servaas A. Morré^{1,2,8}, Andries E. Budding^{5,6}

Affiliations:

- 1: Laboratory of Immunogenetics, Department of Medical Microbiology and Infection Control, VU University Medical Center, Amsterdam, the Netherlands
- 2: Tubascan, spin-off at the Department of Medical Microbiology and Infection Control, VU University Medical Center, Amsterdam, the Netherlands
- 3: Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynaecology, Erasmus University Medical Center, Rotterdam, The Netherlands
- 4: Microbiome Ltd., spin-off at the Department of Medical Microbiology and Infection Control, VU University Medical Center, Amsterdam, The Netherlands
- 5: IS-Diagnostics Ltd., spin-off at Department of Medical Microbiology and Infection Control, VU University Medical Center, Amsterdam, The Netherlands
- 6: Department of Medical Microbiology and Infection Control, VU University Medical Center, Amsterdam, the Netherlands
- 7: Department of Medical Microbiology, Maastricht University Medical Center, Maastricht, the Netherlands
- 8: Institute of Public Health Genomics, Department of Genetics and Cell Biology, Research Institute GROW, Faculty of Health, Medicine & Life Sciences, University of Maastricht, Maastricht, The Netherlands

In preparation

ABSTRACT

16s rRNA gene sequencing is currently the most common way of determining the composition of microbiota. This technique has enabled many new discoveries to be made regarding the relevance of microbiota to the health of the host. However, compared to other diagnostic techniques, 16s rRNA gene sequencing is fairly costly and labor intensive, leaving room for other techniques to improve on these aspects. This study aimed to compare the output of 16s rRNA gene sequencing to the output of IS-pro analysis, using both vaginal swabs and urine samples from 297 women. 16s rRNA gene sequencing and IS-Pro analyses yielded very similar vaginal microbiome profiles, with a median Pearson's R^2 of 0.97. The low bacterial DNA loads of the urine samples prevented adequate analyses using 16s rRNA gene sequencing. However, comparing vaginal and urine samples from individual patients characterized by IS-pro with Pearson's R analysis showed a high correlation (median R^2 = 0.78). This indicates a high level of similarity between 16s rRNA gene sequencing and IS-pro results and between vaginal and urinary microbiota. From this we conclude that 16s rRNA gene sequencing and IS-pro can both be used to determine the microbiota composition. Interestingly, due to having only one amplification step in its pipeline, IS-pro appears to process lower load samples better than can be done through conventional 16s rRNA gene sequencing. Conversely, the quantity and format of output of 16s rRNA gene sequencing is better suited for deeper analysis of the results using currently available bioinformatics tools.

Importance

The microbiota in the vagina has been extensively studied through the use of 16s rRNA gene sequencing. However, compared to other diagnostic techniques, this takes a relatively long time and can be a labor intensive process. In this study we compared 16s rRNA gene sequencing with another technique called IS-pro. We found that, even though IS-pro analysis was comparatively easier and quicker to execute, results of the analysis were comparable to those of the 16s rRNA gene sequencing. Additionally, analyzing the urine samples was only possible with IS-pro, as the process is more capable of handling samples that do not contain large amounts of bacteria. We conclude that IS-pro is a suitable alternative for the profiling of vaginal and urinary microbiota in women.

INTRODUCTION

Bacterial microbiota in humans has received increasing attention over the past decade. Although links between microbiota and host health have been made for a long time, advances such as 16s rRNA gene sequencing have only recently made it possible to properly characterize an individual's microbiome. This has led to many new links between host microbiome and disease(1, 2). Something that is less studied compared to the gut microbiome, but that is more relevant to women's health, is the vaginal microbiome. Commensal microbiota in the vagina control pH levels through the production of lactic acid which is thought to provide a barrier to opportunistic pathogens.

Previous studies have shown that the vagina of healthy women is usually dominantly colonized by a large amount of one out of a limited number of different lactobacilli (3). The four most common of these lactobacilli dominant vaginal microbiome profiles are characterized by either *L. iners*, *L. crispatus*, *L. gasseri*, or *L. jensenii*. However, not all women have a *Lactobacillus* dominant vaginal flora, a condition which is associated with the clinical condition of Bacterial Vaginosis (BV). BV is a disruption of the ecological vaginal balance by overgrowth of typically non-Lactobacillus anaerobic bacteria (4). This results in an alteration of the milieu and composition of the vaginal microbiome (4, 5). BV is the most common vaginal disorder in women and occurs in up to 20% of pregnant women (6). Notable BV related bacteria are *Gardnerella vaginalis*, *Mobiluncus spp.*, and *Atopobium vaginae* (7-9). Symptoms of BV include itchiness and a grey, watery discharge with a fishy odor. However, roughly 50% of the women who have BV are asymptomatic or have less obvious symptoms (10).

For decades the consensus has been that urine is sterile under normal circumstances. However, recently this claim has been disputed, suggesting that a distinct microbiome exists in the urinary tract with potential effects on host health (11). This urinary microbiome shows similarities with that of the vagina, however the amount of bacteria is generally much lower. The bacteria found in the urinary tract can be related to the vaginal flora of the same individual (12).

Microbiome profiling is becoming a highly important tool for diagnosis and prediction for a range of clinical phenomena. Currently, 16s rRNA gene sequencing are seen as the standard way of obtaining microbiome profiles. 16s rRNA gene sequencing, however, is still a relatively

expensive, and labor intensive procedure. As a cheaper and faster alternative, we have also included the IS-pro technique which is based on analysis of the length of the 16S–23S rRNA intergenic spacer (IS) region, which is indicative for bacterial species (13). IS-pro has the additional benefit of being able to process low load samples very efficiently.

In this study we aimed to perform an accurate comparison of microbiome profile outputs produced through 16s rRNA gene sequencing and IS-pro analysis, using vaginal swab and urine samples sequentially taken from women prior to IVF or IVF-ICSI treatment. The resulting data show to what extent 16s rRNA gene sequencing and IS-pro analysis are comparable in their ability to determine the microbiota compositions of women from both high and low load samples. Additionally, it gives insight into the microbiota found in the vagina and urine of women that are about to undergo IVF or IVF-ICSI.

MATERIALS AND METHODS

Sampling

Included in this study were 297 women attending reproductive clinics for IVF or IVF/ICSI treatment. Informed consent forms were obtained from all participants. Midstream urine samples and vaginal swabs were self-collected at one of eight participating reproductive health clinics from June 2015 until March 2016. Urine samples were stored at room temperature for a maximum 2 hours until further processing or storage at 2–8°C. Within 24 hours urine samples were vortexed and concentrated by centrifugation of 10 ml of urine for 10 min at 1500 RCF followed by resuspension in 1 ml of urine which was stored at -20°C. Vaginal swabs were collected by subject at the collection sites and directly placed in 0.5 ml of reduced transport fluid (RTF, Microbiome, Amsterdam, The Netherlands) at 2–8°C for a maximum of 2 hours, after which the swab is stored at -20°C. Samples were transferred on dry ice and stored at -20°C until further processing.

DNA extraction and sample preparation

DNA was extracted from concentrated urine and vaginal swabs suspensions with the Chemagen (Perkin-Elmer, Baesweiler, Germany) automated DNA extraction machine using the buccal swab extraction kit according to the manufacturer's instructions. In short, swab suspensions and urine were thawed and vortexed. 200 µl of sample was incubated with

200 µl Chemagen lysis buffer and 10 µl Proteinase K at 56°C while shaking at 500 rpm. DNA was extracted with the protocol buccal Swab Prefilling. Elution of DNA was in 100µl of Chemagen Elution buffer.

Formation of the library

Sample DNA concentration was measured with the Picogreen dsDNA assay (ThermoFisher, MA, USA). A PCR amplifying the V3/V4 region of the 16S rRNA gene region was performed with individually distinguishable dual index primer sets, which were developed to distinguish low diversity microbiomes, on each sample as has previously been described by Fadrosch *et al.* (2014) (14). The universal primer set 319F/806R, altered to also encode the Illumina sequencing primer and barcode labelling sequences, was used during the PCR. PCR conditions were as follows: 30 seconds at 98°C, then 30 cycles of 10 seconds at 98°C, 15 seconds at 58°C, and 15 seconds at 72°C and a final step of 3 minutes at 72°C.

The amplified DNA was purified with the AMPure XP magnetic bead assay (BeckmanCoulter Genomics, Danvers, MA, USA) quantified as above, recalculated into nM with the formula: “[nM DNA] = DNA concentration (ng/µl) x 1e6 (µl/L) / (Sample fragment size in bp x 656,4 (g/mole))” and equalized to 12 nM. To ensure quality, pooled DNA that did not reach at least 8 nM was not used for 16s rRNA gene sequencing analysis.

16s rRNA gene sequencing 16s rRNA gene sequencing of the pooled samples was performed by the Tumor Genome Analysis Core group of the Department of Pathology at the VU University Medical Center in Amsterdam, The Netherlands with a Miseq tabletop sequencer (Illumina, San Diego, CA, USA).

Sequencing data analysis

Data generated through the 16s rRNA gene sequencing was processed with QIIME to remove primer and index sequences. A minimum Phred quality score threshold of 5 was upheld throughout the processing. Paired end reads with no errors in the barcode matching, a minimum overlap of six nucleotides, and a minimum combined length of 400 nucleotides were assembled to produce identifiable sequences. Operational Taxonomic Units (OTU) were picked with the Usearch method (15)@.. During this process the sequences were sorted based on length and abundance of identical reads, checked for chimeric sequences, and clustered at 97% identity to denoise the data. These OTUs were aligned to the reference

database with the PyNAST method for sequence alignment and subsequently assigned with the RDP classifier method which uses a Naïve Bayes classification. The assignment of OTUs used the database previously described by Srinivasan *et al.* (2012), assigning sequences on a genus to species level (16). The remaining sequences were BLASTed, and included if the sequence in question could be identified at a genus or species level.

Intergenic spacer profiling (IS-pro)

Amplification of 16S-23S rRNA intergenic spacer (IS)-regions was performed with the IS-pro assay (IS-diagnostics, Amsterdam, The Netherlands). IS-pro differentiates bacterial species by the length of the 16S–23S rRNA IS-region with taxonomic classification by phylum-specific fluorescently labeled PCR primers (13). The assay consists of two multiplex PCRs: one PCR contains two different fluorescently-labeled primers: one for the phyla Actinobacteria, Firmicutes, Fusobacteria and Verrucomicrobia and a second color for the phylum Bacteroidetes. A separate PCR is performed for the phylum Proteobacteria. The assay was performed according to the protocol provided by the manufacturer. Amplifications were carried out on a GeneAmp PCR system 9700 (Applied Biosystems, Foster City, CA). After PCR, 5 µl of PCR product was mixed with 20 µl formamide and 0.5 µl Mapmaker 1500 ROX-labeled size marker (BioVentures, Murfreesboro, TN, USA). DNA fragment analysis was performed on an ABI Prism 3500 Genetic Analyzer (Applied Biosystems). Species were assigned to peaks by using a database compiled of IS-pro fragments obtained from in-silico and in vitro IS-pro PCRs of known urine and vagina associated bacterial species. An internal amplification control (IAC) was used to control the PCR reaction for inhibition. A sample passed the quality control when the IAC signal was present in sufficient amount (3 of 5 IAC peaks >500 Relative Fluorescence Units (RFU)) or when a sufficiently high bacterial signal was present (at least one bacterial peak >20.000 RFU).

Data analysis

Alpha diversity of the microbiome per sample was measured by calculating the Shannon diversity index of individual samples. Relative abundance of microbiome per sample was used to perform a correlation clustering of all sample profiles according to the UPGMA method. Relative abundance for 16s rRNA gene sequencing data was calculated as a percentage of reads from total reads; for IS-pro, relative abundance is given as fluorescence intensity per peak as a percentage of total fluorescence. This data was then used to identify the major clusters making up the datasets. Pearson's R linear regression was used

to compare abundance of species between samples. For Pearson's R calculations only species that were available in both the 16s rRNA gene sequencing and IS-pro databases were included.

RESULTS

16s rRNA gene sequencing quality control

After sequencing of the vaginal samples, all reads were monitored for quality control purposes. A total of 294 (of 297) vaginal samples produced sequences matching the quality criteria. Sequencing of the DNA resulted in 17,947,706 reads. 8,374,321 reads passed quality control and were assigned to a taxon. Vaginal samples yielded a median of 9,661 reads per sample. During OTU calling, 75 species and 22 genera were assigned to the samples. An average of 29 (Stdev: 21,7) species or genera were assigned per vaginal sample. The two included sequencing controls yielded on average 9 reads after processing. Resulting taxonomic distributions can also be seen in Table S1. Pooled DNA from the urine samples did not meet the required 8nM of DNA for proper 16s rRNA gene sequencing analysis and was therefore not included in the 16s rRNA gene sequencing analysis.

16s rRNA gene sequencing results of vaginal samples

The heatmap in Figure 1 shows the microbiome profiles of the vaginal samples displaying relative abundance of bacterial species, clustered based on cosine-correlation. The Shannon diversity index of the samples shows a clear increase in diversity of the microbiome in profiles that are not clearly dominated by a single species, in most cases *L. crispatus* and *L. iners*.

Notable clusters identified through hierarchical clustering of the vaginal microbiome profiles included a *L. crispatus* dominant cluster including 132 samples, a *L. gasseri* dominant cluster including 17 samples, a *L. iners* dominant cluster including 74 samples, a diverse microbiome profile including 38 samples, and a *L. jensenii* dominant cluster including 22 samples. Eleven vaginal samples could not be ascribed to any cluster according to our clustering. These samples were characterized by a number of non-lactobacillus dominant bacteria, e.g. *Leptotrichia* or *Prevotella* that were not defining for the observed clusters.

IS-pro results of vaginal samples

All of the 297 vaginal samples yielded results that passed the quality control as described in the Materials and Methods. The heatmap in Figure 2 shows the microbiome profiles of these vaginal samples characterized with IS-pro, displaying relative abundance of bacterial species clustered based on cosine-correlation. All IS-pro taxonomic distributions can be found in Table S2. Notable clusters identified through hierarchical clustering of the vaginal microbiome profiles were generally similar to clusters found with 16s rRNA gene sequencing for characterization and included a *L. crispatus* dominant cluster with 133 samples, a *L. gasseri* dominant cluster with 14 samples, a *L. iners* dominant cluster with 129 samples, a diverse microbiome profile with 12 samples, and a *L. jensenii* dominant cluster with seven samples. Two vaginal samples could not be ascribed to any specific cluster. These samples were characterized by a number of non-lactobacillus dominant bacteria. e.g. *Leptotrichia* or *Prevotella*.

IS-pro results of urine samples

The heatmap in Figure 3 shows the microbiome profiles of the urine samples characterized with IS-pro, displaying relative abundance of bacterial species, and clustered based on cosine-correlation. Notable clusters identified through hierarchical clustering include: a *L. crispatus* dominant cluster with 75 samples, a *L. gasseri* dominant cluster with 12 samples, a *L. iners* dominant cluster with 125 samples, a *L. jensenii* dominant cluster with six samples, and an *E. coli* dominant cluster with eight samples. A diverse microbiome was found in four samples. 26 urine samples could not be ascribed to any specific cluster. Internal amplification control was sufficient in 41 samples that did not provide any other signal and were therefore determined to be negative.

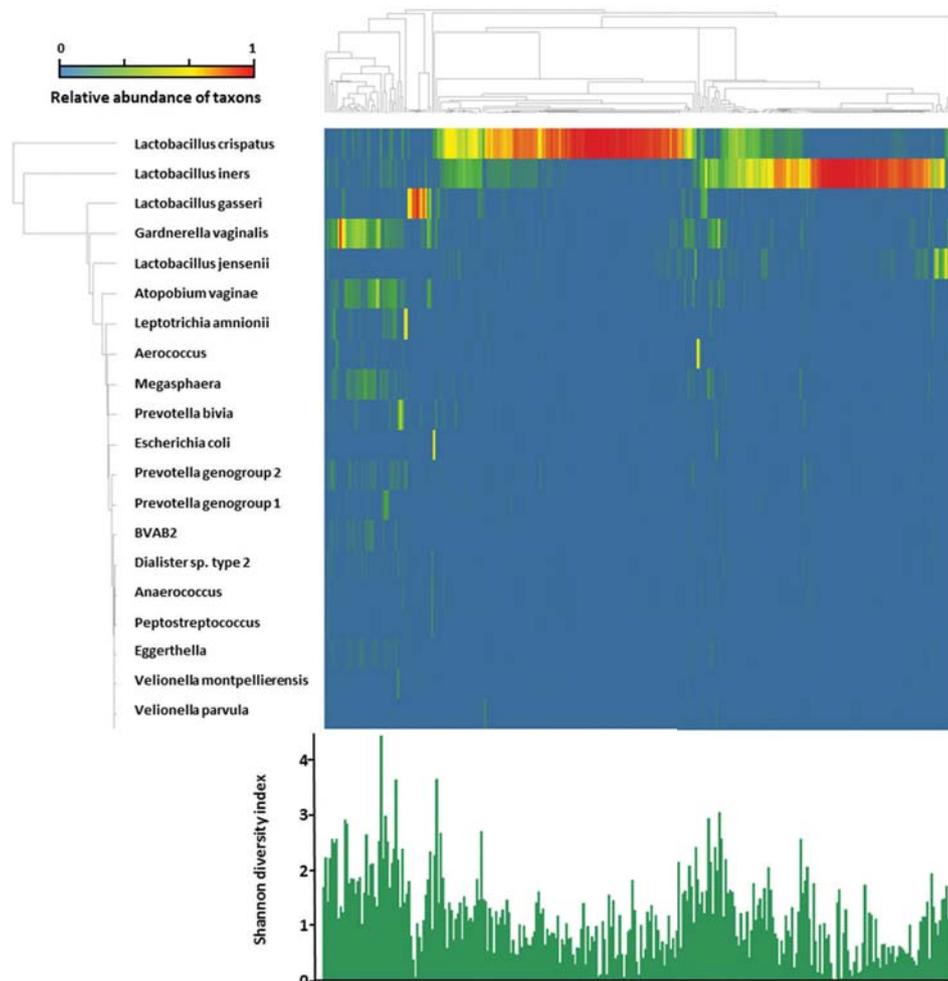


Figure 1: Heatmap of relative microbiome abundance found in vaginal samples obtained from 294 women through 16s rRNA gene sequencing. Column correlation clustering was performed with the UPGMA method to cluster microbiome profiles according to similarity. Row hierarchy clustering was performed to identify and order the most prominent taxa related to the microbiome profiles. Shown in the figure are the 20 most abundant taxa found in this correlation. The alpha diversity is shown in the bar graph using the Shannon diversity index in a sample order correlating with the above heatmap profiles.

CHAPTER 8

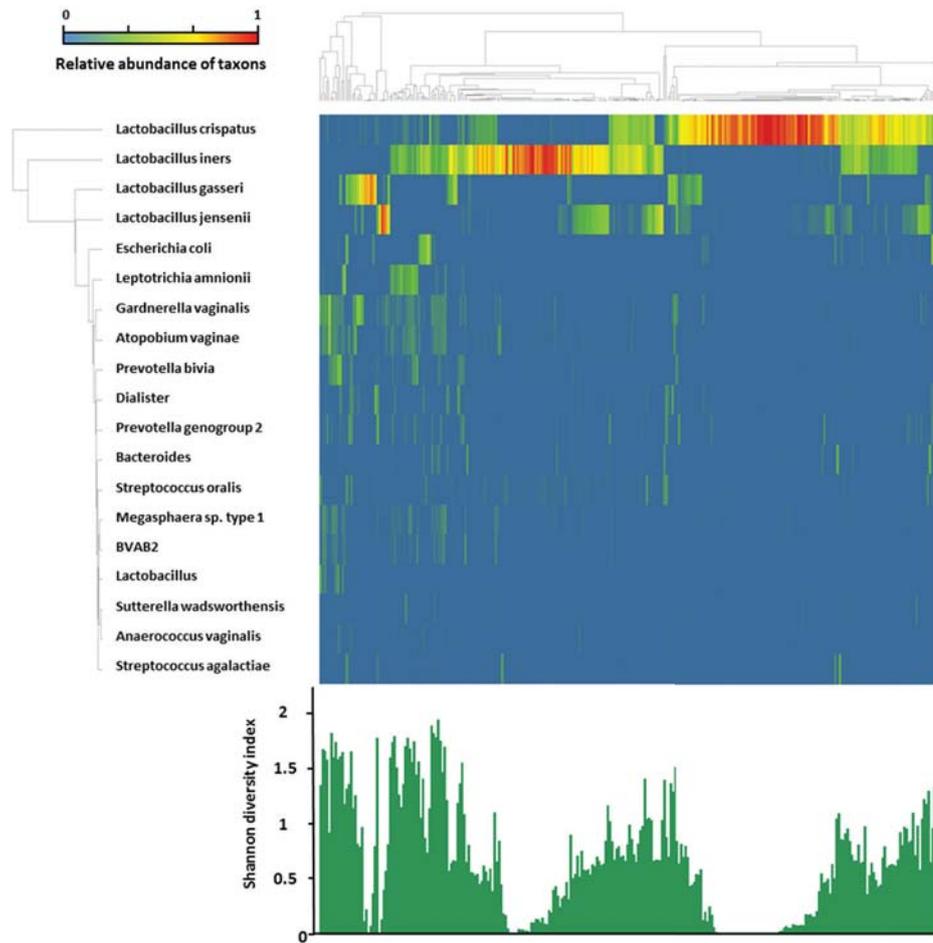


Figure 2: Heatmap of relative microbiome abundance found in vaginal samples obtained from 297 women with IS-pro. Column correlation clustering was performed with the UPGMA method to cluster microbiome profiles according to similarity. Row hierarchy clustering was performed to identify and order the most prominent taxa related to the microbiome profiles. Shown in the figure are the 19 most abundant taxa found in this correlation. The alpha diversity is shown in the bar graph using the Shannon diversity index in a sample order correlating with the above heatmap profiles.

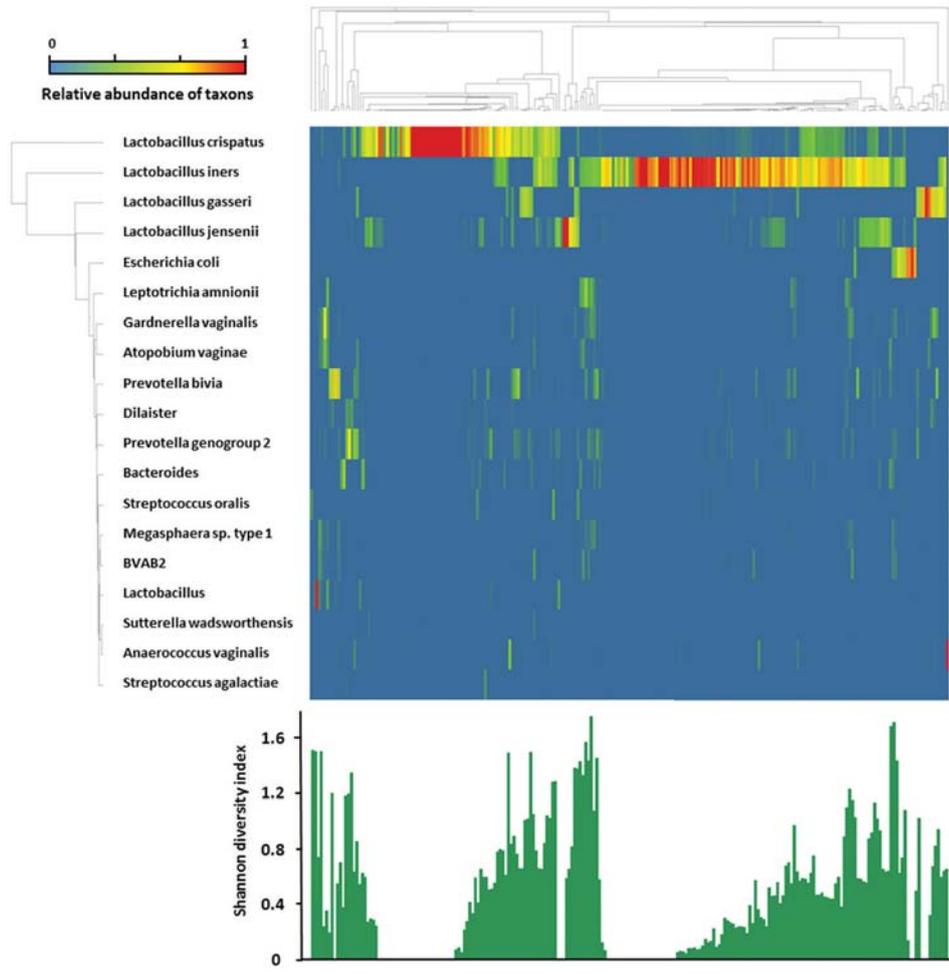


Figure 3: Heatmap of relative microbiome abundance found in urine samples obtained from 256 women attending reproductive clinics for IVF or IVF/ICSI treatment through IS-pro. Column correlation clustering was performed with the UPGMA method to cluster microbiome profiles according to similarity. Row hierarchy clustering was performed to identify and order the most prominent taxa related to the microbiome profiles. Shown in the figure are the 19 most abundant taxa found in this correlation. The alpha diversity is shown in the bar graph using the Shannon diversity index in a sample order correlating with the above heatmap profiles.

Comparison of 16s rRNA gene sequencing and IS-pro profiles of vaginal samples

Table 1 shows a cross table depicting distribution of vaginal sample profile clusters between 16s rRNA gene sequencing and IS-pro. The two methods yielded almost completely consistent *L. crispatus* cluster assignments. Sixty-eight samples were assigned to the *L. iners* cluster by both methods. To statistically determine the comparability of the cluster compositions of 16s rRNA gene sequencing and IS-pro, we further compared the results of 16s rRNA gene sequencing and IS-pro by calculating Pearson's R correlation in paired samples from the same patient. This comparison showed a high correlation of the IS-pro and 16s rRNA gene sequencing results, with a median R^2 of 0.97 (Figure 4A and 4B).

Table 1: Distribution of vaginal sample cluster profiles between 16s rRNA gene sequencing results and IS-pro results, respectively. Only samples successfully analyzed by both techniques are shown.

		16s rRNA gene sequencing vaginal profiles					
		Divers	<i>L. crispatus</i>	<i>L. gasseri</i>	<i>L. iners</i>	<i>L. jensenii</i>	Other
IS-pro	Divers	6	2	2	-	-	-
	<i>L. crispatus</i>	4	113	3	6	3	3
	<i>L. gasseri</i>	4	1	8	-	-	1
	<i>L. iners</i>	21	16	4	68	13	7
	<i>L. jensenii</i>	1	-	-	-	6	-
	Other	2	-	-	-	-	-

Comparison of vaginal and urine profiles in IS-pro

Next, we compared urine and vaginal profiles from individual patients using the IS-pro generated profiles. The Pearson's calculations yielded a R squared of 0.78 (Figure 4C), indicative of a strong correlation between vaginal data and urine data. Furthermore, Figure 5 shows the total (5A) and relative (5B) abundance of bacteria for both samples of each subject on a horizontal line. In general, dominant species, such as *L. crispatus* (blue bars), *L.gasseri* (pink bars), and *L. iners* (yellow bars) are shared between vaginal and urine profiles of the same individual. Urine profiles of *L. crispatus* dominated individuals appear to contain a higher diversity of other bacteria than those of *L. iners* dominated individuals. *E. coli* has a high abundance in a number of urine samples.

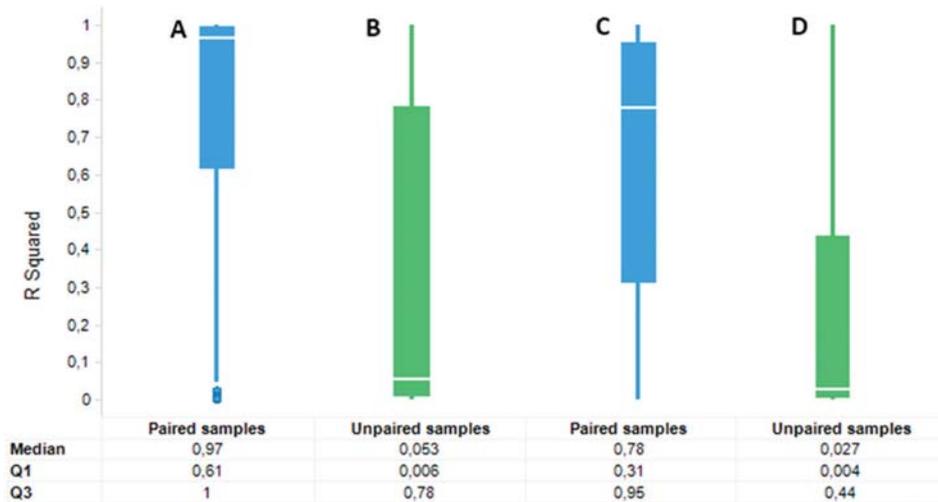


Figure 4: Outcomes of Pearson's R correlation where blue bars represent outcomes from analyses based on paired sample numbers, and green bars represent the same analyses where samples were not paired based on sample number. **A & B** Boxplot featuring R squared values of IS-pro vaginal sample outcomes correlated to those of 16s rRNA gene sequencing vaginal sample outcomes when samples are paired based on sample number (**A**) vs no pairing (**B**). **C & D** Boxplot featuring R squared values of IS-pro vaginal sample outcomes correlated to IS-pro urine sample outcomes when samples are paired based on sample number (**C**) vs no pairing (**D**). (Q1= 1st quartile, Q3 = 3rd quartile)

CHAPTER 8

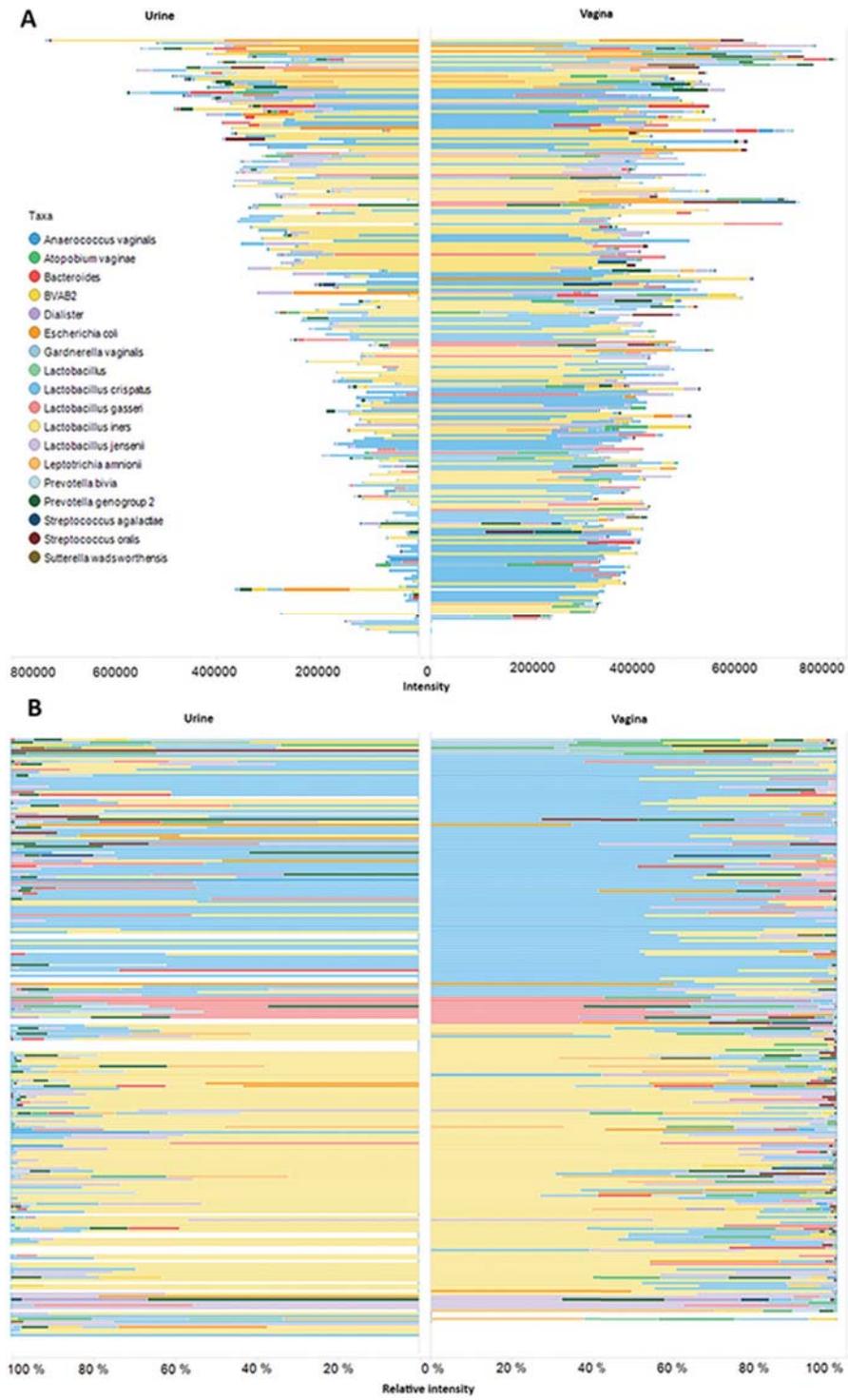


Figure 5: Tornado plots where each horizontal line represents one subject with urine profiles on the left and vaginal profiles on the right of the vertical axis. White bars indicate samples that failed to yield data during analysis of one sample type. **(A)** IS-pro intensity distribution per subject depicting found taxa per color, and **(B)** the relative intensity of species signals found with IS-pro per subject. Included in the figure is a key clarifying which color depicts which taxa and showing the most relevant taxa found through IS-pro.

DISCUSSION

16s rRNA gene sequencing is currently the most common way of determining the composition of microbiota. However, compared to other diagnostic techniques, 16s rRNA gene sequencing is fairly costly and labor intensive, creating space for other techniques. In this study we show that 16s rRNA gene sequencing and IS-pro analysis produce comparable outcomes when profiling microbiota from vaginal swabs and urine samples. When sufficient DNA was present for IS-pro analysis, we found that microbiome profiles of urine samples were highly comparable to vaginal swabs, albeit with lower total loads.

The focus of our study was the comparison of two techniques used for microbiome profiling. 16s rRNA gene sequencing is currently seen as the gold standard for the profiling of microbiome. This is despite a lack of reproducibility between laboratory settings which is described in literature to be caused by, among other factors, the use of different DNA extraction procedures, library creation protocols, and/or sequencing equipment (17). In this study we show that vaginal microbiome profiling using the IS-pro technique creates outcomes highly comparable to those of the 16s rRNA gene sequencing. In addition, we show that the IS-pro is more capable of dealing with lower load samples, as urine samples could only be analyzed using the IS-pro due to load limitations that are set for 16s rRNA gene sequencing. This highlights the IS-pro technique as a test to be used in lieu of 16s rRNA gene sequencing for microbiome profiling, as was also shown in a previous study (13).

Even though the analyses were highly comparable, if we expand on the comparison of 16s rRNA gene sequencing and IS-pro techniques, we still observed a number of differences between them. Although the data produced by either 16s rRNA gene sequencing or IS-pro produced similar profiles there are a number of practical differences that are significant depending on the needs and limitations of the user. Microbiome profiling through 16s rRNA gene sequencing allows for the use of a vast array of bio-informatics tools that are the result

CHAPTER 8

of many years of development as demand grew. This makes it more likely that complex custom demands can be applied on the user data. In contrast, IS-pro was developed with the goals of cost-effectiveness and simplicity in mind. This means that both the laboratory processing of the sample and the data-analysis were completed faster than for the 16s rRNA gene sequencing analysis, and at reduced cost.

In the results of the vaginal profiling using both 16s rRNA gene sequencing and IS-pro, profiles were found to be highly similar. However, when looking at the clustering data, there are still a number of differences in the amount of samples per cluster. This is especially apparent in the more diverse clusters. These differences are generally caused by small differences in abundance of species in profiles per technique. No formal criteria were used for clustering samples, but clusters were formed by performing a UPMGA clustering on a cosine correlation matrix. Therefore, small differences between samples may lead to different clustering outcomes.

The similarity of the vaginal and urinary samples can perhaps be attributed to contamination of urine samples by passage through the urethra and vulva, however this is a highly controversial topic. There are a number of fairly recent studies that do suggest there is a distinct urinary microbiota (18-22). The difficulty of obtaining samples from the urinary tract that are certain not to be contaminated by other microbiota as proof of the existence or non-existence of these microbiota has hindered clarifying research on this topic. In a study by Wolfe et al. suprapubic aspirates were obtained from patients specifically to avoid downstream contamination in the urinary tract (23). Even through use of this technique, a number of bacteria related to the vaginal microbiota were still found in the urine. These results were later confirmed in a study by Jacobs et al. who found that most bacteria in the urine were either *Lactobacillus* spp. or *Gardnerella vaginalis* (24). These findings, in combination with the results from our study, indicate that the microbiota found in the urine samples of this study are likely due to spillover of bacteria from the vagina and urethra.

Looking back at this study, some strengths and limitations should be discussed. The parallel analyses of 16s rRNA gene sequencing and IS-pro give a unique opportunity to compare the techniques with regards to output quality which are factors that could not be properly assessed if only one technique had been used. Furthermore, this serves as a check to see if any one technique obtains unexpected results compared to the other.

As a first limitation, the samples collected for this study were collected from women who were about to undergo an IVF or IVF-ICSI treatment due to subfertility. Although vaginal microbiome composition distributions found in this study showed similar distributions as other studies into the vaginal microbiome, it is still possible that this sampling group instigated a sampling bias for certain profiles. Second, the databases used for the 16s rRNA gene sequencing and IS-pro data processing did not completely overlap and were specifically composed for the vaginal microbiome. This may have resulted in missing bacteria species in either technique. Finally, it is likely that potential PCR bias plays a more significant role in the 16s rRNA gene sequencing procedure than in the IS-pro analysis as the former entails two PCR reactions. The strict 16s rRNA gene sequencing input criteria aiming at negating the effect of the resulting amplification competition is the reason why urine samples analyzed through 16s rRNA gene sequencing could not be included in this study.

The low loads of bacterial DNA prevented us from using 16s rRNA gene sequencing to profile urine samples in this study. In fact, the median DNA concentration in urine samples was 108 ng/ml, compared to 7007 ng/ml in the vaginal samples (data not shown). Besides the fact that this made the urine samples impossible to use in 16s rRNA gene sequencing analysis, we can also interpret from this data that the bacterial abundance in the urinary tract is very low.

In conclusion 16s rRNA gene sequencing and IS-pro analysis produce highly comparable results when analyzing microbiota collected with vaginal swabs. We also conclude that IS-pro is more suitable for analysis of microbiota collected through urine sampling, as the low bacterial load does not allow for analysis with standard 16s rRNA gene sequencing protocols. IS-pro analysis has the potential to increase speed and reduce costs of these analyses while maintaining the quality of the profiling, hopefully allowing for more research to take place into the vaginal microbiome.

DECLARATIONS

Acknowledgements

The authors would like to acknowledge the efforts of dr. R van Houdt and C. van der Veer from the department of Medical Microbiology at the VUmc, Amsterdam in the setup of the 16s rRNA gene sequencing protocol. Additionally, the authors greatly appreciate the collaborating centers: Radboud University Medical Center, Erasmus University Medical Center, Isala kliniek Zwolle, VU University Medical Center, Sint Elisabeth Ziekenhuis, VivaNeo Medisch Centrum Kinderwens, University Medical Center Utrecht, and University Medical Center Maastricht for the timely sample inclusions.

Author's contributions

MS performed the 16S rRNA gene sequencing and analyses. RK performed the sample collection. MPB performed the DNA extraction and assisted with the 16S rRNA gene sequencing data analysis. LP performed IS-pro analysis and data processing. JSEL, SAM, PHMS and AEB conceived and designed the experiments. Additionally, AEB performed the majority of IS-pro data analysis. MS, MPB and AEB drafted the manuscript and all authors read and approved its final version.

Competing interests

The authors would like to state a number of competing interests. The authors P.H.M. Savelkoul and A.E. Budding are co-owners of IS-diagnostics Ltd., which is the company that developed the IS-pro technique. The author R. Koedooder reports that she is an employee at ARTPred B.V. during her PhD at Erasmus Medical Centre. Joop S.E. Laven reports consultancy fees from Titus Health Care.

Ethics approval and consent to participate

The medical ethics testing committee Erasmus MC has approved the ethicality of the study under reference MEC-2014-455. All participants were informed of the study contents and signed an informed consent form before inclusion into the study.

Availability of data and materials

The datasets generated and/or analysed during the current study are available in the SRA database repository, <https://www.ncbi.nlm.nih.gov/sra/SRP133380> .

The database that supports the IS-pro findings of this study are available from IS-Diagnostics Ltd. but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of IS-Diagnostics Ltd.

Supplemental Files

Due to their size, the supplemental files of this article are not shown in this thesis. They are available upon request, and will also be available at the time this article is published.

REFERENCES

1. Joossens M, Huys G, Cnockaert M, De Preter V, Verbeke K, Rutgeerts P, et al. Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. *Gut*. 2011;60(5):631-7.
2. Willing BP, Finlay BB. Gut microbiology: fitting into the intestinal neighbourhood. *Curr Biol*. 2009;19(11):R457-9.
3. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, et al. Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A*. 2011;108 Suppl 1:4680-7.
4. Nugent RP, Krohn MA, Hillier SL. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *J Clin Microbiol*. 1991;29(2):297-301.
5. Srinivasan S, Morgan MT, Fiedler TL, Djukovic D, Hoffman NG, Raftery D, et al. Metabolic signatures of bacterial vaginosis. *MBio*. 2015;6(2).
6. Donati L, Di Vico A, Nucci M, Quagliozzi L, Spagnuolo T, Labianca A, et al. Vaginal microbial flora and outcome of pregnancy. *Arch Gynecol Obstet*. 2010;281(4):589-600.
7. Donders GG, Vereecken A, Bosmans E, Dekeersmaecker A, Salembier G, Spitz B. Definition of a type of abnormal vaginal flora that is distinct from bacterial vaginosis: aerobic vaginitis. *BJOG*. 2002;109(1):34-43.
8. Hernandez-Rodriguez C, Romero-Gonzalez R, Albani-Campanario M, Figueroa-Damian R, Meraz-Cruz N, Hernandez-Guerrero C. Vaginal microbiota of healthy pregnant Mexican women is constituted by four *Lactobacillus* species and several vaginosis-associated bacteria. *Infect Dis Obstet Gynecol*. 2011;2011:851485.
9. Lamont RF, Sobel JD, Akins RA, Hassan SS, Chaiworapongsa T, Kusanovic JP, et al. The vaginal microbiome: new information about genital tract flora using molecular based techniques. *BJOG*. 2011;118(5):533-49.
10. Ralph SG, Rutherford AJ, Wilson JD. Influence of bacterial vaginosis on conception and miscarriage in the first trimester: cohort study. *BMJ*. 1999;319(7204):220-3.
11. Lewis DA, Brown R, Williams J, White P, Jacobson SK, Marchesi JR, et al. The human urinary microbiome; bacterial DNA in voided urine of asymptomatic adults. *Front Cell Infect Microbiol*. 2013;3:41.
12. Gottschick C, Deng ZL, Vital M, Masur C, Abels C, Pieper DH, et al. The urinary microbiota of men and women and its changes in women during bacterial vaginosis and antibiotic treatment. *Microbiome*. 2017;5(1):99.
13. Budding AE, Grasman ME, Lin F, Bogaards JA, Soelton-Kaersenhout DJ, Vandenbroucke-Grauls CM, et al. IS-pro: high-throughput molecular fingerprinting of the intestinal microbiota. *FASEB J*. 2010;24(11):4556-64.
14. Fadrosh DW, Ma B, Gajer P, Sengamalay N, Ott S, Brotman RM, et al. An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome*. 2014;2(1):6.
15. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. 2010;26(19):2460-1.
16. Srinivasan S, Hoffman NG, Morgan MT, Matsen FA, Fiedler TL, Hall RW, et al. Bacterial communities in women with bacterial vaginosis: high resolution phylogenetic analyses reveal relationships of microbiota to clinical criteria. *PLoS One*. 2012;7(6):e37818.
17. Lozupone CA, Stombaugh J, Gonzalez A, Ackermann G, Wendel D, Vazquez-Baeza Y, et al. Meta-analyses of studies of the human microbiota. *Genome Res*. 2013;23(10):1704-14.
18. Price TK, Dune T, Hilt EE, Thomas-White KJ, Kliethermes S, Brincat C, et al. The Clinical Urine Culture: Enhanced Techniques Improve Detection of Clinically Relevant Microorganisms. *J Clin Microbiol*. 2016;54(5):1216-22.
19. Thomas-White K, Taege S, Johansen D, Hilt EE, Brincat C, Mueller ER, et al. Bladder And Vaginal Microbiomes Have A Corresponding Shift Following Estrogen Treatment In Post-Menopausal Women. *The FASEB Journal*. 2017;31(1 Supplement):940.4-4.

20. Thomas-White KJ, Hilt EE, Fok C, Pearce MM, Mueller ER, Kliethermes S, et al. Incontinence medication response relates to the female urinary microbiota. *Int Urogynecol J*. 2016;27(5):723-33.
21. Pearce MM, Hilt EE, Rosenfeld AB, Zilliox MJ, Thomas-White K, Fok C, et al. The female urinary microbiome: a comparison of women with and without urgency urinary incontinence. *MBio*. 2014;5(4):e01283-14.
22. Pearce MM, Zilliox MJ, Rosenfeld AB, Thomas-White KJ, Richter HE, Nager CW, et al. The female urinary microbiome in urgency urinary incontinence. *Am J Obstet Gynecol*. 2015;213(3):347 e1-11.
23. Wolfe AJ, Toh E, Shibata N, Rong R, Kenton K, Fitzgerald M, et al. Evidence of uncultivated bacteria in the adult female bladder. *J Clin Microbiol*. 2012;50(4):1376-83.
24. Jacobs KM, Thomas-White KJ, Hilt EE, Wolfe AJ, Waters TP. Microorganisms Identified in the Maternal Bladder: Discovery of the Maternal Bladder Microbiota. *AJP Rep*. 2017;7(3):e188-e96.