Limitations of 16S rRNA Gene Sequencing to Characterize Lactobacillus Species in the Upper Genital Tract

Jessica L. O'Callaghan

Dana Willner
*William & Mary, dlwillner@wm.edu*

Melissa Buttini

Flavia Huygens

Elise S. Pelzer

Follow this and additional works at: [https://scholarworks.wm.edu/aspubs](https://scholarworks.wm.edu/aspubs)

Recommended Citation

O'Callaghan, Jessica L.; Willner, Dana; Buttini, Melissa; Huygens, Flavia; and Pelzer, Elise S., Limitations of 16S rRNA Gene Sequencing to Characterize Lactobacillus Species in the Upper Genital Tract (2021). *Frontiers in Cell and Developmental Biology*, 9. [https://doi.org/10.3389/fcell.2021.641921](https://doi.org/10.3389/fcell.2021.641921)
Limitations of 16S rRNA Gene Sequencing to Characterize Lactobacillus Species in the Upper Genital Tract

Jessica L. O'Callaghan1,2, Dana Willner3,4, Melissa Buttini5, Flavia Huygens1,2 and Elise S. Pelzer1,2*

1 Faculty of Health, Centre for Immunology and Infection Control, Queensland University of Technology, Brisbane, QLD, Australia, 2 Faculty of Health, School of Biomedical Sciences, Queensland University of Technology, Brisbane, QLD, Australia, 3 Australian Centre for Ecogenomics, University of Queensland, St Lucia, QLD, Australia, 4 Department of Computer Science, College of William & Mary, Williamsburg, VA, United States, 5 The Wesley Hospital, Auchenflower, QLD, Australia

The endometrial cavity is an upper genital tract site previously thought as sterile, however, advances in culture-independent, next-generation sequencing technology have revealed that this low-biomass site harbors a rich microbial community which includes multiple Lactobacillus species. These bacteria are considered to be the most abundant non-pathogenic genital tract commensals. Next-generation sequencing of the female lower genital tract has revealed significant variation amongst microbial community composition with respect to Lactobacillus sp. in samples collected from healthy women and women with urogenital conditions. The aim of this study was to evaluate our ability to characterize members of the genital tract microbial community to species-level taxonomy using variable regions of the 16S rRNA gene. Samples were interrogated for the presence of microbial DNA using next-generation sequencing technology that targets the V5–V8 regions of the 16S rRNA gene and compared to speciation using qPCR. We also performed re-analysis of published data using alternate variable regions of the 16S rRNA gene. In this analysis, we explore next-generation sequencing of clinical genital tract isolates as a method for high throughput identification to species-level of key Lactobacillus sp. Data revealed that characterization of genital tract taxa is hindered by a lack of a consensus protocol and 16S rRNA gene region target allowing comparison between studies.

Keywords: Lactobacillus, genital tract, 16S rRNA, next-generation sequencing, low-biomass

INTRODUCTION

Molecular microbiology techniques have changed our ability to identify microbial communities, revolutionizing the way we assess female genital tract microorganisms. In cultivation-dependent studies, greater than 95% of the vaginal microbiota in healthy women was classified as lactobacilli (Klebanoff et al., 1991). The advent of cultivation-independent technology platforms provided evidence to suggest that in up to two-thirds of healthy women, the lactobacilli were co-aggregated with a diverse group of microbial community members, and in some cases did not dominate
Samples were collected from women undergoing operative hysterectomy and laparoscopy. Paired endometrial curettages and endocervical swabs were obtained aseptically from 56 women at the time of surgery. The vagina and perineum were disinfected with povidone iodine at the time of anesthetic induction. An endocervical swab (Amies Agar Transport Swab, Copan) was collected prior to cervical dilatation, followed by disinfection of the ectocervix prior to collection of endometrial curettages and then collection of a second endocervical swab. Specimens were refrigerated at the time of collection and processed within 1 h.

DNA was extracted from 1 mL aliquots of solutions removed from the transport swabs. Endometrial curettages were suspended in 1.5 mL of sterile physiological saline, from which DNA was extracted using the QiAMP Mini DNA extraction kit (Qiagen, Australia) as per the manufacturer’s instructions with the addition of a preliminary enzymatic lysis step and final elution in 50 µL of sterile water.

**MATERIALS AND METHODS**

A basic flowchart of the methods including in this study can be found as Supplementary Figure 1.

**Patient Cohort, Sample Collection, and Genomic DNA Preparation**

Samples were collected from women undergoing operative hysterectomy and laparoscopy. Paired endometrial curettages and endocervical swabs were obtained aseptically from 56 women at the time of surgery. The vagina and perineum were disinfected with povidone iodine at the time of anesthetic induction. An endocervical swab (Amies Agar Transport Swab, Copan) was collected prior to cervical dilatation, followed by disinfection of the ectocervix prior to collection of endometrial curettages and then collection of a second endocervical swab. Specimens were refrigerated at the time of collection and processed within 1 h.

DNA was extracted from 1 mL aliquots of solutions removed from the transport swabs. Endometrial curettages were suspended in 1.5 mL of sterile physiological saline, from which DNA was taken for DNA extraction. The remaining samples were stored frozen at −80°C. DNA was extracted using the QIAMP Mini DNA extraction kit (Qiagen, Australia) as per the manufacturer’s instructions with the addition of a preliminary enzymatic lysis step and final elution in 50 µL of sterile water.

**Abbreviations**: ATCC, American type culture collection; DGC, dysmenorrhea progestin effect endocervix; DGE, dysmenorrhea progestin effect endometrium; DNA, deoxyribonucleic acid; DPC, dysmenorrhea proliferative endocervix; DPE, dysmenorrhea proliferative endometrium; DSC, dysmenorrhea secretory endocervix; DSE, dysmenorrhea secretory endometrium; HRM, high resolution melt; MGC, menstruation progestin effect endocervix; MGE, menstruation progestin effect endometrium; MPE, menstruation proliferative endocervix; MPE, menstruation proliferative endometrium; MSC, menstruation secretory endocervix; MSE, menstruation secretory endometrium; OUT, operational taxonomic unit; rRNA ribosomal ribonucleic acid; VIC, virgo intacta.

**Table 1** | Sample abbreviations and clinical numbers.

| Sample abbreviations | Clinical patients (n)^a^ | Full sample name |
|----------------------|--------------------------|------------------|
| DGC                  | 8                        | Dysmenorrhea progestin effect endocervix |
| DGE                  | 13                       | Dysmenorrhea progestin effect endometrium |
| DPC                  | 11                       | Dysmenorrhea proliferative endocervix |
| DPE                  | 4                        | Dysmenorrhea proliferative endometrium |
| DSC                  | 7                        | Menorrhagia progestin effect endocervix |
| MGC                  | 4                        | Menorrhagia progestin effect endometrium |
| MGE                  | 7                        | Menorrhagia proliferative endocervix |
| MPC                  | 7                        | Menorrhagia proliferative endometrium |
| MPE                  | 10                       | Menorrhagia secretory endocervix |
| MSC                  | 3                        | Menorrhagia secretory endometrium |
| MCE                  | 3                        | Menorrhagia secretory endometrium |
| VIC                  | 3                        | Virgo intacta |

^a Each patient contributed both an endocervix sample and endometrium sample.

Clinical sample cohorts were constructed as previously described (Pelzer et al., 2018a). Briefly, endometrial curettages were fixed and stained with hematoxylin and eosin for menstrual staging and pathology assessment by the pathology department at Sullivan Nicolaides Pathology. The phase of the menstrual cycle at the time of surgery was reported for each patient. The patient samples from all 12 sample types (Table 1) were combined for the below analysis. Genomic DNA was extracted from individual samples prior to pooling an equal volume of DNA from each individual sample.

**Details of Ethical Approval**

All patients recruited for this study provided written informed consent. Ethical approval was obtained from the review boards of UnitingCare Health (approval number 2013.04.75), Human Research Ethics Committee and Queensland University of Technology Human Ethics Committee (approval number 1400000686). All methods in this study were performed in accordance with the relevant guidelines and regulations as dictated by these ethics boards.

**Next-Generation Sequencing**

The 16S rRNA gene PCR assay was performed using the previously published primers, 803F (5’-ATTAGATACCCTGGTAGTC-3’) and 1392R (5’-ACGGGCCTGTGTTGTC-3’) and PCR cycling conditions (Willner et al., 2014). Fusion primers with 454 adaptor sequences were ligated to the 803F and 1392R primers to amplify the V5 and V8 regions of the 16S rRNA gene (Willner et al., 2014). PCR reactions were performed as previously described (Pelzer et al., 2018a). The five frequently encountered genital tract Lactobacillus sp. gene sequences were aligned to each other using sequences download from the SILVA database \(^1\) to determine the degree of variation within the V5-V8 regions of the 16S rRNA gene. The annealing site of the sequencing primers is marked on the alignment (Figure 1A).

\(^1\)https://www.arb-silva.de/
**Lactobacillus sp.-Specific Quantitative Real-Time PCR**

Quantitative real-time PCR assays were performed using previously published primer pairs (Table 2) and cycling conditions. A standard curve was generated using *L. gasseri* ATCC strain 19992. Primer annealing was confirmed using species-specific alignment of the five *Lactobacillus* sp. interrogated in this study (Figure 1B).

**Taxonomic Classification**

Sequence clustering and operational taxonomic unit (OTU) selection was performed using a modified version of CD-HIT-OTU-454 which does not remove singleton clusters (Liu et al., 2011). Taxonomy was assigned to representative sequences by comparison to the latest build of the Greengenes database using BLAST, and OTU tables were constructed from the output using a custom Perl script (McDonald et al., 2012).

**Lactobacillus Phylogenetic Trees**

Full-length 16S rRNA sequences for *Lactobacillus* spp. (accession numbers: AB680529.1, AB690249.1, AB668940.1, AB008203.1, AB425941.1, AB008206.1, AF243169.1, AF243167.1, CP018809.253324, CP018809.1516019, CP018809.1347636, CP018809.500868, AB547146.1, AB932527.1, AB008209.1, HZ485829.7, LG085736.7 LF134126.7, 2 https://greengenes.secondgenome.com/ 3 https://blast.ncbi.nlm.nih.gov/Blast.cgi
TABLE 2 | 16S rRNA Lactobacillus species-specific primers (Pelzer et al., 2018a).

| Lactobacillus species | Primer name | Primer sequence |
|-----------------------|-------------|-----------------|
| L. acidophilus         | LAA-1       | 5′CATCCAGTGCAAACCTAAGAG3′ |
|                       | LAA-2       | 5′GATGCGCTTGCCTCTGCA3′ |
| L. crispatus           | LcrisF      | 5′AGCGACGCGAATAACAGATTAC3′ |
|                       | LcrisR      | 5′AGCTGATCATGCGATCTGCT3′ |
| L. gasseri             | LgassF      | 5′AGCGACGCTGGCTAGATGAATTTG3′ |
|                       | LgassR      | 5′TCCTTTAACTCTAGACATGC3′ |
| L. jensenii            | LjensF      | 5′AAGTCGAGCGAGCTTGCTATAGA3′ |
|                       | LjensR      | 5′CTCTTCTATGCGAAAGTAGC3′ |
| L. iners               | InersF      | 5′GTCGCCTTQAAGATCGG3′ |
|                       | InersR      | 5′ACAGTTGATAAGCATC3′ |

LG104504.7), Pediococcus pentosaceus (accession numbers: AB018215.1 and AB362987.1), and Bacillus subtilis (accession numbers: AP012496.9810 and AP012496.30276) were downloaded from the SILVA database using the web interface. Sequences were aligned using ClustalW (Thompson et al., 1994) with the default settings. MEGA7 (Kumar et al., 2016) was used to generate the best-known maximum likelihood (ML) tree using a Jukes-Cantor model and 1000 bootstrapping iterations. The ML tree was visualized and edited in FigTree⁴ and Adobe Illustrator (Figure 1).

To generate the phylogenetic trees for the specific regions, the same full-length 16S rRNA sequences from above were imported into Geneious⁵ along with two Escherichia coli sequences downloaded from SILVA (see text footnote 1) (accession number: AB045730.1 and AB045731.1). Sequences were aligned using standard Geneious alignment and trimmed to include the specific variable regions (V1–V2, V1–V3, V3–V4, V3–V4, V4, and V5–V8) using E. coli sequences. Trimmed sequences were then aligned using ClustalW with default settings and imported into MEGA7, and trees were constructed and edited in the same way as above (Figure 2).

**Hierarchical Clustering**

A dissimilarity matrix was generated based on the relative abundances of Lactobacillus spp. in the pyrosequenced and qPCR analyzed samples using the vegdist function in the vegan package in R⁶ with the Bray-Curtis dissimilarity metric (Oksanen et al., 2009). Hierarchical clustering was performed using the hclust function in R with “average” linkage (UPGMA) (R Core Team, 2014). Clustering and relative abundances were visualized in a heatmap with associated dendrogram using the heatmap.2 function from the R package ggplots (Wickham, 2016).

**Re-analysis of Published Endometrial 16S rRNA Sequencing Datasets**

Four datasets (Table 3) were re-analyzed using the standard QIIME2⁷ (Bolyen et al., 2019) data analysis pipeline to determine Lactobacillus abundance using a standard full-length taxonomic

---

⁴http://tree.bio.ed.ac.uk

⁵https://www.geneious.com/

⁶https://www.r-project.org/

⁷http://qiime.org/

FIGURE 2 | 16S rRNA gene phylogeny for key genital tract lactobacilli (A) Full length (B) V5–V8 region (C) V1–V2 region (D) V3–V5 region (E) V1–V3 region (F) V3–V4 region (G) V4 – The different color for the top clade is due to the mixed L. acidophilus and L. crispatus sequences.
TABLE 3 | Comparison of publicly available 16S rRNA endometrial datasets used for the sequencing re-analysis.

| Dataset | Sample cohort | Sequencing | Primers | Analysis | Total | Total Lactobacillus abundance |
|---------|---------------|------------|---------|----------|-------|-----------------------------|
| PRJNA329174 (Moreno et al., 2016) | Endometrial fluid from fertile women at distinct time points before and during pregnancy | 454 pyrosequencing | V3–V5<sub>a</sub> | Identified 191 OTUs and stratified samples as either Lactobacillus dominant or non-Lactobacillus dominant. The non-Lactobacillus dominant endometrium was correlated with decreases in implantation, ongoing pregnancy, and live birth rates | 2126 | 248 11.66% |
| PRJNA546137 (Hernandes et al., 2020) | Vaginal fluid, healthy endometrium and (endometriosis—deep lesions only) | Illumina single end | V3–V4<sub>b</sub> | No significant differences between control and endometriosis patients but did identify that deep endometriotic lesions are correlated with a Lactobacillus lacking environment | 348 | 32 9.19% |
| PRJEB18626 (Wee et al., 2018) | Vaginal, cervical, and endometrial samples of control and women with a history of infertility | Illumina paired end 300 | V1–V3<sub>c</sub> | Communities with dominant L. crispatus and L. iners were identified | 55 | 0% |
| PRJNA543861 (Winters et al., 2019) | Mid endometrial samples from women undergoing a hysterectomy and compared with other parts of the reproductive tract | Illumina 500 paired end | V4<sub>d</sub> | Concluded that the endometrium is not Lactobacillus dominated but is comprised of Acinetobacter, Pseudomonas, Cloacibacterium species | 316 | 2 0.63% |

<sup>a</sup>(F342–926) > Forward CCTACGGGAGGCAGCAG > Reverse CGGTCAATTYMTTTTRAGT.
<sup>b</sup>(27–534) > Forward CCTACGGGAGGCAGCAG > Reverse GGACTACHVGGGTWTCTAAT.
<sup>c</sup>(27–534) > Forward AGAGTTTGATYMTGGCTTA > Reverse TACGGGGCTGCTGGCA.
<sup>d</sup>(515–806) > Forward GTGCCAGCMGCCGCGGTAA > Reverse GGACTACHVGGGTWTCTAAT.

RESULTS

16S rRNA Amplicon Sequencing Resolution of Genital Tract Lactobacillus sp. Operational Taxonomic Units (OTUs) to Genus and Species

In order to determine the Lactobacillus species present in upper genital tract (UGT) clinical samples (Table 1), 16S rRNA gene sequencing was performed on the V5-V8 variable region. Ten operational taxonomic units (OTUs) identified in the genital tract were attributable to Lactobacillus sp. [Lactobacillus sp. genus level (n = 2), L. crispatus (n = 2), L. iners (n = 3), L. intestinalis (n = 1), L. jensenii (n = 1), and L. vaginalis (n = 1)]. The majority of Lactobacillus sp. OTUs (8/10) were resolved to the genus and species level exploiting the V5–V8 regions of the 16S rRNA gene (Pelzer et al., 2018a).

Lactobacillus Species-Specific Quantitative Real-Time PCR Assay Comparison to 16S rRNA Amplicon Sequencing Output

To confirm the predicted species taxonomic assignment from the 16S rRNA amplicon sequencing output, species-specific quantitative real-time PCR assays were used to confirm the species.
identity and relative abundance of only five of the most prevalent species in the UGT (L. crispatus, L. gasseri, L. iners, L. jensenii, and L. acidophilus) in comparison to a standard curve.

The abundance of the five most prevalent species (L. crispatus, L. gasseri, L. iners, L. jensenii, and L. acidophilus) were then compared between qPCR and relative abundances as determined by amplicon sequencing (Figure 3). L. crispatus and L. iners predominated the Lactobacillus community in most samples when characterized by 16S rRNA amplicon sequencing. All samples displayed similar abundance profiles with enriched lower abundance species including L. jensenii, L. gasseri, L. iners, and L. acidophilus detected by qPCR (Figure 3 and Supplementary Table 2). As shown in Figure 3, L. crispatus and L. iners predominated the microbial communities in samples analyzed in this study and samples were more likely to cluster based on Lactobacillus community predominance, than the patient history (dysmenorrhea or menorrhagia), the anatomical site of collection (endometrium or cervix), or the analysis technique (qPCR or NGS).

**Lactobacillus Phylogeny**

In order to understand if the use of different variable regions of the 16S rRNA gene results in changes to taxonomic classification, phylogenetic trees were constructed using the full-length 16S rRNA gene sequences of the Lactobacillus sp. described in this study. These trees indicated that L. acidophilus and L. crispatus appear as sister groups derived from a single node for all primer pairs except V4, where the two species appear as a single group of descendnt taxa (Figure 2).
Primer Pair Comparison Using Published Endometrial 16S rRNA Sequencing Data

To determine if alternative regions of the 16S rRNA gene lead to changes for Lactobacillus species prediction, publicly available datasets with different primer pairs were used for Lactobacillus species comparisons, primer biasing analysis, and host DNA amplification prediction. All four published datasets used variable region primer pairs targeting different regions of the 16S rRNA gene (Table 3). The total predicted abundance of Lactobacillus sp. varied from 11.66% (V3–V5) and 9.19% (V3–V4) to 0.63% (V4) and 0% (V1–V3). Of the four datasets, three were Illumina technologies and two of these were paired-end studies (300 base pair and 500 base pair). However, despite the differences in experimental design, datasets were independently classified into amplicon sequence variants (ASVs) and then used for taxonomic analysis to allow comparison of results. OTUs were used for the original data analysis in our earlier publications characterizing the upper genital tract microbiome. Therefore, we retained the data in OTU format for consistency and included it in the original form in this manuscript. ASVs were used in the re-analysis of the four published datasets because that is now the accepted method of taxonomic classification. The two methods of analysis were not ever directly compared because OTUs are a combination of similar sequences and ASVs are not. None of the Lactobacillus sp. of interest in this study were speciated using the standard QIIME2 taxonomic classifier. Sequence coverage differences between primer pairs for the simulated control dataset was statistically significant (p = 2e-16).

Potential primer sequencing bias was identified for all primers in the study, including the V5-V8 primers used by our group (Table 4 and Figure 4). All primer pairs inaccurately reported the abundance of Lactobacillus sp. in the lower genital tract. All primer pairs except for those targeting the V4 region had predicted human genome amplification (data not shown).

DISCUSSION

This study examined the bacterial composition of the communities in the upper genital tract of women, reporting changes in the bacterial community composition of lactobacilli, revealing that the primer selection may significantly bias the predicted microbial community profile. Importantly, 16S rRNA primer pairs used to characterize the low-biomass upper genital tract should be selected to most accurately amplify targeted lactobacilli. Sequencing of specific variable regions of the 16S rRNA gene improves the discriminatory power for speciation of dominant genital tract lactobacilli (Graspeuntenner et al., 2018). Further, there exists a compromise in primer selection for correct reporting of taxonomic classification and species abundance and low host genome amplification for low-biomass sites. Current studies may select primers that have low host genome amplification, such as the V4 region, but at the cost of reduced sensitivity in differentiating species such as L. acidophilus and L. crispatus (evidenced in Figure 2). There also exists a growing body of evidence highlighting the cross-amplification of host DNA using 16S rRNA primers (Walker et al., 2020).

### Table 4

| Database | Primer pairs total # of reads that were attributed to each species from the 20000 and the percentage of total reads that were from each species | Species | Total # | % V1–V2 | % V3–V5 | % V3–V4 | % V1–V3 | % V4 | % V5–V8 |
|----------|-------------------------------------------------------------------------------------------------------------------------------------|--------|---------|---------|---------|---------|---------|---------|---------|
| L. gasseri | 635 | 64,70 | 32.95 | 46,355 | 23.11 | 5,580 | 28.07 | 5,761 | 28.76 |
| L. iners | 76 | 1,821 | 9.15 | 1,821 | 9.15 | 1,283 | 6.41 | 1,185 | 5.92 |
| L. crispatus | 311 | 5,031 | 25.15 | 5,031 | 25.15 | 5,341 | 26.71 | 5,753 | 28.76 |
| L. jensenii | 82 | 1,918 | 9.59 | 1,918 | 9.59 | 1,423 | 7.11 | 1,275 | 6.37 |
| L. acidophilus | 267 | 4,760 | 23.80 | 4,760 | 23.80 | 6,153 | 30.78 | 5,936 | 29.68 |
| Total | 1,371 | 20,000 | 100% | 20,000 | 100% | 20,000 | 100% | 20,000 | 100% |

| Forward | Reverse |
|---------|---------|
| APCAGAGTTTGATCNTGGCTCAG | TGCTGCCTCCCGTAGGAGT |
| CCTACGGGAGGCAGCAG | CCGTCAATTYMTTTRAGT |
| CCTACGGGRSGCAGCAG | GGACTACHVGGGTWTCTAAT |
| AGAGTTTGATYMTGGCTCA | TTACCGCGGCTGCTGGCA |
| GTGCCAGCMGCCGCGGTAA | GGACTACHVGGGTWTCTAAT |
Sequencing technologies frequently only report the presence of lactobacilli at the genus level. Studies exploiting some regions of the 16S rRNA gene fail to discriminate lactobacilli beyond higher order taxonomic classification due to limited sequence variation (Brown et al., 2018; Wee et al., 2018; Bernabeu et al., 2019; Zhang et al., 2019; Al-Memar et al., 2020; Tu et al., 2020). Therefore, some studies have reported that lactobacilli in the lower genital tract as a genera are: positively correlated with healthy pregnancy outcomes including successful implantation and delivery at term; and form abundant community members in cases of infertility, adverse pregnancy outcomes including recurrent implantation failure and preterm birth, and cancer (Moreno et al., 2016; Chen et al., 2017; Miles et al., 2017). For example, *L. iners* in the lower genital tract reportedly does not protect against preterm birth and is frequently reported as an abundant community member in women with bacterial vaginosis suggesting that the presence of any *Lactobacillus* species is not indicative of a health milieu (Madhivanan et al., 2014; Petricevic et al., 2014). In addition, our data suggest that detection of *L. iners* by 16S rRNA amplicon sequencing may represent an overestimation of this species. Further, significant differences between lactobacilli in term compared to preterm deliveries have not been reported for all studies (Romero et al., 2014; Amabebe and Anumba, 2018).

Examination of the genital tract as a continuum concluded that sampling high-biomass lower genital tract sites including the vagina or cervix identified distinct microbial community profiles compared to more diverse, but low-biomass upper genital tract sites including the endometrial cavity, fallopian tubes, or peritoneal cavity (Chen et al., 2017; Pelzer et al., 2018a,b), perhaps highlighting the need for more discriminatory analyses to tease out significant taxa associated with the genital tract sites in health and disease.

*L. crispatus* and *L. iners* were the most abundant lactobacilli in the samples tested by 16S rRNA amplicon sequencing in our study, which is consistent with the previous studies that we reanalyzed as part of this project. We observed the same shortfalls commonly associated with lower order taxonomic classification and low-biomass samples. Similar results can be observed in molecular studies characterizing the female genital tract when multiple variable regions of the 16S rRNA gene were sequenced (Fettweis et al., 2012; Madhivanan et al., 2014; Miles et al., 2017; Graspeuntn, 2018). Van Der Pol et al. (2019) reported that the choice of 16S rRNA reference sequence database and sample sequence clustering parameters are equally as important as the choice of variable region for amplification characterizing microbial community members to lower orders.
There is no doubt that sequencing the conserved 16S rRNA gene has improved our understanding of extant biodiversity in human microbial communities and is critical for understanding the impact of low-abundance community members on health and disease. However, there is no consensus best practice for microbiome studies of the female genital tract, and significant variability exists between sample collection and storage methods, DNA extraction, universal primer selection, and sequencing platform and data analysis software (Pollock et al., 2018). Characterization of microbial communities using the 16S rRNA gene have been hampered by inherent differences generated in community profiles when sequencing different hypervariable regions, short read lengths, and taxonomic classification difficulties due to limited resolution for closely related species (Poretsky et al., 2014). Sequencing technologies have been used to interrogate the genital tract microbial community in reproductive-age women but most fail to resolve the isolates to the species level. The V4 region is favored because amplification does not result in non-specific host DNA amplification, however, as demonstrated by phylogenetic analyses, the ability to confidently resolve lower order taxonomic classification for lactobacilli is a trade-off in genital tract analyses. However, Lactobacillales is a diverse and heterogeneous family with some members genetically very closely related, but lacking a clear monophyletic origin, impeding the ability to accurately identify lower order taxonomy in studies exploiting conserved genes such as the 16S rRNA gene (Heilig et al., 2002).

Consequently, more recent efforts have focused on sequencing multiple variable regions of the gene with amalgamation of all data into a single profile (Fuks et al., 2018). The most recent research focuses on removing bias associated with sequencing component variable regions by using full-length gene sequencing (Callahan et al., 2016). The need to characterize the full-length 16S rRNA gene is further required as exhibited by the change in L. gasseri, L. jensenii, and L. iners clustering when comparing the full-length gene to specific variable regions. The ability to confidently assign the correct identity to key genital tract lactobacilli is hampered by primer selection. Collectively our research confirms what other studies have shown, that health and disease (Figure 4) may depend on species and strain-level differences for prominent community members at a given anatomical niche (Kraal et al., 2014). It is clear that additional discriminatory power is required to resolve lower order taxonomic classifications using current sequencing methods. This report confirms that speciation of key genital tract Lactobacillus sp., capable of modulating reproductive health is possible when the appropriate region of the 16S rRNA gene is interrogated.

CONCLUSION

Studies characterizing microbial communities in the female genital tract report inconsistent results when assessing dysbiosis as a cause of reproductive pathology. This paper provides further evidence for the impact of primer selection on evaluating the biological significance of shifts in community taxa in the endometrial cavity. Careful experimental design should include a comparative analysis of microbial community profiling data generated by interrogation of multiple variable regions to the 16S rRNA gene or the full-length gene to ensure that species abundance and diversity are accurately reflected.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the UnitingCare Health Human Research Ethics Committee and Queensland University of Technology Human Research Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JO'C and DW designed and completed bioinformatics analyses, contributed to the analysis and interpretation of the data, and contributed to the writing of the manuscript. MB conceived and designed the project, performed collection of clinical specimens, and contributed to the writing of the manuscript. FH designed and completed the qPCR experiments, contributed to the analysis and interpretation of the qPCR data, and contributed to the writing of the manuscript. EP conceived and designed the project, completed tissue processing, DNA extraction, and 16S PCR experiments, contributed to the analysis and interpretation of the data, and drafted significant parts of the work. All authors contributed to the article and approved the submitted version.

FUNDING

Funding for this project was awarded by the Wesley Research Institute (Grant No. 2011-12). The funding body played no role in conducting the research or preparing the manuscript.

ACKNOWLEDGMENTS

We thank Wesley Hospital theater staff who facilitated the collection of the genital tract samples. We acknowledge the Australian Centre for Ecogenomics for 454 pyrosequencing and Professor Philip Hugenholz and Dr. Fiona May. This work was performed in the John and Wendy Thorsen Women's Health Laboratory.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.641921/full#supplementary-material
REFERENCES

Al-Memar, M., Bobdiwala, S., Foure, H., Mannino, R., Lee, Y., Smith, A., et al. (2020). The association between vaginal bacterial composition and miscarriage: a nested case–control study. *BIOG* 127, 264–274.

Amababe, E., and Anumba, D. O. C. (2018). The vaginal microenvironment: the physiologic role of Lactobacilli. *Front. Med. (Lausanne)* 5:181. doi: 10.3389/fmed.2018.00181

Angly, F. E., Willner, D., Rohwer, F., Hugenholtz, P., and Tyson, G. W. (2012). Grinder: a versatile amplicon and shotgun sequence simulator. *Nucleic Acids Res.* 40, e94–e94.

Barnabeu, A., Lledo, B., Diaz, MaC, Lozano, F. M., Ruiz, V., Fuentes, A., et al. (2019). Effect of the vaginal microbiome on the pregnancy rate in women receiving assisted reproductive treatment. *J. Assist. Reprod. Genet.* 36, 2111–2119. doi: 10.1007/s10815-019-01564-0

Blostein, F., Gelaye, B., Sanchez, S. E., Williams, M. A., and Foxman, B. (2020). Vaginal microbiome diversity and preterm birth: results of a nested case–control study in Peru. *Ann. Epidemiol.* 41, 28–34. doi: 10.1016/j.annepidem.2019.11.004

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghazli, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857.

Brown, R. G., Al-Memar, M., Marchesi, J. R., Lee, Y. S., Smith, A., Chan, D., et al. (2019). Establishment of vaginal microbiota composition in early pregnancy and its association with subsequent preterm prelabour rupture of the fetal membranes. *Transl. Res.* 207, 30–43. doi: 10.1016/j.trsl.2018.12.005

Brown, R. G., Marchesi, J. R., Lee, Y. S., Smith, A., Lehne, B., Kindinger, L. M., et al. (2018). Vaginal dysbiosis increases risk of preterm fetal membrane rupture, neonatal sepsis and is exacerbated by erythromycin. *BMJ Med.* 169. doi: 10.1186/s12196-017-0999-x

Callahan, B. J., Sankaran, K., Fukuyama, J. A., McMurdie, P. J., and Holmes, S. P. (2016). Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. *F1000Res* 5:1492. doi: 10.12688/f1000research.89862.1

Campisciano, G., Florian, F., D’Eustacchio, A., Stanković, D., Ricci, G., Seta, F. D., et al. (2017). Subclinical alteration of the cervical–vaginal microbiome in women with idiopathic infertility. *J. Cell. Physiol.* 232, 1681–1688. doi: 10.1002/jcp.25806

Chen, C., Song, X., Wei, W., Zhong, H., Dai, J., Lan, Z., et al. (2017). The microbiota continuum along the female reproductive tract and its relation to uterine-related diseases. *Nat. Commun.* 8:875.

Fettweis, J. M., Serrano, M. G., Sheth, N. U., Mayer, C. M., Glascock, A. L., Brooks, J. P., et al. (2012). Species-level classification of the vaginal microbiome. *Genomics* 13:S17. doi: 10.1186/1471-2164-13-S8-S17

Pelzer, E. S., Willner, D., Buttini, M., and Huygens, F. (2018a). A role for the endometrial microbiome in dysfunctional menstrual bleeding. *Antonie Van Leeuwenhoek* 111, 933–943. doi: 10.1007/s10482-017-0999-6

Pelzer, E. S., Willner, D., Huygens, F., and Fischbach, M. A., et al. (2014). Characterisation of the endometrial microbiota has an effect on implantation success or failure. *Am. J. Obstet. Gynecol.* 215, 684–703. doi: 10.1016/j.ajog.2016.09.075

Nasidouss, D., Forney, L. J., Schneider, G. M., Gliniewicz, K., France, M., Boester, S. P. (2016). Bioconductor workflow for microbiome data analysis: from raw Nucleic Acids Research (NAR) data to validated computational tools. *Mol. Biol. Evol.* 33, 1870ñ1874. doi: 10.1093/molbev/msw054

Pelzer, E. S., Willner, D., Buttini, M., and Huygens, F. (2018b). Fallopian tube microbiota: evidence beyond DNA. *Future Microbiol.* 13, 1353ñ1361. doi: 10.2217/fmb-2018-0118

Petričević, L., Domig, K. J., Nierscher, F. J., Sandhofer, M. J., Fidesser, J., Petričević, L., Domig, K. J., Nierscher, F. J., Sandhofer, M. J., Fidesser, J., et al. (2014). The Female Genital Tract. *Sci Rep* [Internet]. Available online at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5759028/ (accessed Nov 10, 2020).

Oksanen, J., Kindt, R., Legendre, P., Haras, B., Simpson, G., Solymos, P., et al. (2009). The Vegan Package. Available online at: https://www.researchgate.net/publication/228339454_The_Vegan_Package

Petričević, L., Domig, K. J., Nierscher, F. J., Sandhofer, M. J., Fidesser, J., Krondorfer, L., et al. (2014). Characterisation of the vaginal Lactobacillus microbiota associated with preterm delivery. *Sci. Rep.* 5:1536.

Pollock, J., Glendingin, L., Wisedchawan, T., and Watson, M. (2018). The madness of microbiome: attempting to find consensus “Best Practice” for 16S microbiome studies. *Liu S-J, editor. Appl. Environ. Microbiol.* 84:e2627–17.

Poretzky, R., Rodríguez-R, L. M., Luo, C., Tsenteml, D., and Konstantinidis, K. T. (2014). Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. *PLoS One* 9:e93827. doi: 10.1371/journal.pone.0093827

Qu, W., Shen, Z., Zhao, D., Yang, Y., and Zhang, C. (2009). MFEprimer: multiple factor evaluation of the specificity of PCR primers. *Bioinformatics*. 25, 276–278. doi: 10.1093/bioinformatics/btn614

R Core Team, (2014). *R: A Language and Environment for Statistical Computing*. *Internet*. Vienna: R Foundation for Statistical Computing.

Ravel, J., Gajer, P., Abdo, Z., Schneider, G. M., Koenig, S. S. K., McCulle, S. L., et al. (2011). Vaginal microbiome of reproductive-age women. *Proc. Natl. Acad. Sci. USA* 108(Supplement 1), 4680–4687.

Romero, R., Hassan, S. S., Gajer, P., Tarca, A. L., Fadrosh, D. W., Bieda, J., et al. (2014). The vaginal microbiota of pregnant women who subsequently have spontaneous preterm labor and delivery and those with a normal delivery at term. *Microbiome* 2:18.

Seo, S. S., Arakijaray, S., Kim, M. K., Oh, H. Y., Kwon, M., Kong, J. S., et al. (2017). High Prevalence of Leptotrichia amnionii, Atopobium Vaginae, Sneathia sanguinengens, and Factor I Microbes and Association of Spontaneous Abortion...
among Korean Women. Biomed Res Int [Internet]. Available online at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5745682/ (accessed Nov 10, 2020).

Tu, Y., Zheng, G., Ding, G., Wu, Y., Xi, J., Ge, Y., et al. (2020). Comparative Analysis of Lower Genital Tract Microbiome Between PCOS and Healthy Women. Front Physiol [Internet]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7506141/ (accessed Nov 10, 2020).

Van Der Pol, W. J., Kumar, R., Morrow, C. D., Blanchard, E. E., Taylor, C. M., Martin, D. H., et al. (2019). In silico and experimental evaluation of primer sets for species-level resolution of the vaginal microbiota using 16S ribosomal RNA gene sequencing. J. Infect. Dis. 219, 305ñ314. doi: 10.1093/infdis/jiy508

Walker, S. P., Barrett, M., Hogan, G., Flores Bueso, Y., Claesson, M. J., and Tangney, M. (2020). Non-specific amplification of human DNA is a major challenge for 16S rRNA gene sequence analysis. Sci. Rep. 10:16356.

Wee, B. A., Thomas, M., Sweeney, E. L., Frentiu, F. D., Samios, M., Ravel, J., et al. (2018). A retrospective pilot study to determine whether the reproductive tract microbiota differs between women with a history of infertility and fertile women. Aust. N. Zool. J. Obstet. Gynaecol. 58, 341ñ348. doi: 10.1111/ajo.12754

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis[Internet]. Available online at: https://ggplot2.tidyverse.org/ (accessed October 24, 2020).

Copyright © 2021 O’Callaghan, Willner, Buttini, Huygens and Pelzer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.