

# SCIENTIFIC REPORTS



OPEN

## Does the endometrial cavity have a molecular microbial signature?

Andrew D. Winters<sup>1,2,3</sup>, Roberto Romero<sup>3,4,5,6,7</sup>, Maria Teresa Gervasi<sup>8</sup>, Nardhy Gomez-Lopez<sup>1,2,3,9</sup>, Maria Rosa Tran<sup>8</sup>, Valeria Garcia-Flores<sup>3,9</sup>, Percy Pacora<sup>3,9</sup>, Eunjung Jung<sup>3,9</sup>, Sonia S. Hassan<sup>2,3,9,10</sup>, Chaur-Dong Hsu<sup>2,3,9,10</sup> & Kevin R. Theis<sup>1,2,3</sup>

Recent molecular studies concluded that the endometrium has a resident microbiota dominated by *Lactobacillus* spp. and is therefore similar to that of the vagina. These findings were largely derived from endometrial samples obtained through a transcervical catheter and thus prone to contamination. Herein, we investigated the molecular microbial profiles of mid-endometrial samples obtained through hysterectomy and compared them with those of the cervix, vagina, rectum, oral cavity, and controls for background DNA contamination. Microbial profiles were examined through 16S rRNA gene qPCR and sequencing. Universal bacterial qPCR of total 16S rDNA revealed a bacterial load exceeding that of background DNA controls in the endometrium of 60% (15/25) of the study subjects. Bacterial profiles of the endometrium differed from those of the oral cavity, rectum, vagina, and background DNA controls, but not of the cervix. The bacterial profiles of the endometrium and cervix were dominated by *Acinetobacter*, *Pseudomonas*, *Cloacibacterium*, and Comamonadaceae. Both 16S rRNA gene sequencing and *Lactobacillus* species-specific (*L. iners* & *L. crispatus*) qPCR showed that *Lactobacillus* was rare in the endometrium. In conclusion, if there is a microbiota in the middle endometrium, it is not dominated by *Lactobacillus* as was previously concluded, yet further investigation using culture and microscopy is necessary.

The endometrium has traditionally been considered sterile in non-pregnant women, in pregnant women, and in the puerperium based upon cultivation studies<sup>1–16</sup>. Indeed, infection of the endometrium or the presence of high concentrations of microbial products (e.g. endotoxin) has been associated with implantation failure<sup>17,18</sup>, spontaneous abortion<sup>19</sup>, recurrent miscarriage<sup>20</sup>, and spontaneous preterm birth<sup>21,22</sup>.

It is difficult to envision that the mucosa of the endometrial cavity could be continuously exposed to microorganisms present in the lower genital tract, as well as to sperm that can potentially carry microorganisms into the uterus, and yet remain free of bacterial colonization<sup>22</sup>. Recent studies using molecular techniques suggest that there may be a resident microbiota (i.e., the assemblage of microorganisms present in a defined environment<sup>23</sup>) in the human endometrium<sup>24–37</sup>, and the potential relevance of these microbes to women's reproductive health, especially fertilization and normal pregnancy outcomes, is being considered<sup>26,28,38</sup>. For instance, *Lactobacillus*-dominance (>90% relative abundance) of the endometrial microbiota has been associated with implantation success<sup>28,34</sup> and live birth rates in women undergoing *in vitro* fertilization (IVF)<sup>28</sup>. It has therefore been suggested that an increase in relative abundance of *Lactobacillus* species to >90% in women with a non-*Lactobacillus*-dominated endometrial microbiota might promote implantation success among infertile patients<sup>34</sup>.

<sup>1</sup>Department of Biochemistry, Microbiology, and Immunology, Wayne State University School of Medicine, Detroit, Michigan, USA. <sup>2</sup>Perinatal Research Initiative in Maternal, Perinatal and Child Health, Wayne State University School of Medicine, Detroit, Michigan, USA. <sup>3</sup>Perinatology Research Branch, Division of Obstetrics and Maternal-Fetal Medicine, Division of Intramural Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, U.S. Department of Health and Human Services, Bethesda, MD, and Detroit, MI, Detroit, Michigan, USA. <sup>4</sup>Department of Obstetrics and Gynecology, University of Michigan, Ann Arbor, Michigan, USA. <sup>5</sup>Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, Michigan, USA. <sup>6</sup>Center for Molecular Medicine and Genetics, Wayne State University, Detroit, Michigan, USA. <sup>7</sup>Detroit Medical Center, Detroit, Michigan, USA. <sup>8</sup>Department of Women's and Children's Health, University Hospital of Padua, Padua, Italy. <sup>9</sup>Department of Obstetrics and Gynecology, Wayne State University School of Medicine, Detroit, Michigan, USA. <sup>10</sup>Department of Physiology, Wayne State University School of Medicine, Detroit, Michigan, USA. Correspondence and requests for materials should be addressed to R.R. (email: [prbchiefstaff@med.wayne.edu](mailto:prbchiefstaff@med.wayne.edu)) and K.R.T. (email: [ktheis@med.wayne.edu](mailto:ktheis@med.wayne.edu))

Received: 23 April 2019

Accepted: 21 June 2019

Published online: 09 July 2019

Nevertheless, a fundamental question remains as to whether an endometrial microbiota truly exists<sup>39</sup>. The potential limitations of previous reports are two-fold. First, most of the previous molecular surveys investigating an endometrial microbiota analyzed samples that had been collected transcervically (11/15 studies in Supplementary Table 1); this sampling approach is prone to contamination of the endometrium with microbes and/or microbial molecular signals from the vagina and cervix. Indeed, a culture-based investigation of endometrial bacteria demonstrated an increased rate of bacterial isolation from samples obtained transcervically than transabdominally<sup>10</sup>. As a result, the widely reported *Lactobacillus*-dominance of the endometrial microbiota may be due to the influence of contamination with vaginal *Lactobacillus* species during sampling. A second reason further investigation is necessary is that, if an endometrial microbiota truly exists, then it is present at very low biomass and thus its molecular characterization is susceptible to influences of background DNA contamination from extraction kits and PCR and sequencing reagents (collectively referred to as the “kitome”)<sup>40–44</sup>. As a result, contaminating DNA may constitute a considerable portion, if not all, of the observed molecular microbial signatures within the endometrium. It is therefore necessary that molecular investigations of an endometrial microbiota incorporate technical controls for potential sources of background DNA contamination and provide detailed descriptions of the microbial profiles of these controls when characterizing the endometrial microbiota. To date, the majority of sequence-based molecular surveys investigating an endometrial microbiota either have not incorporated technical controls or have not provided detailed descriptions of the microbial profiles of these controls (10/14 sequence-based studies in Supplementary Table 1).

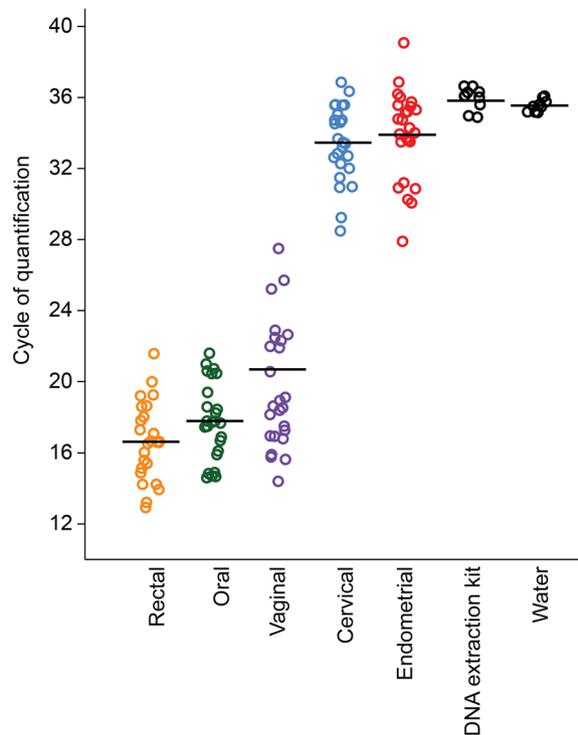
Given the potential of both the sample collection method and background DNA contamination to shape characterizations of endometrial microbiota profiles, the existence of a resident endometrial microbiota, and its structure if indeed present, remains unknown<sup>39</sup>. The objective of this study was to use 16S rRNA gene sequencing, universal bacterial 16S rDNA qPCR, and *Lactobacillus*-targeted qPCR to characterize the bacterial profiles of endometrial samples from non-pregnant women who underwent transabdominal hysterectomy. The study further determined if the endometrial microbiota is distinct from that of other body sites by comparing the bacterial profiles of the endometrium to those of the cervix, vagina, rectum, and oral cavity of the women, as well as to technical controls for potential background DNA contamination.

## Materials and Methods

**Study population.** This was a cross-sectional study of 25 women with a median age of 45 years (IQR: 41–49.5) who underwent a hysterectomy, primarily for fibroids (23/25), at the Azienda Ospedaliera di Padova, Italy. The two women without fibroids underwent hysterectomy for endometrial hyperplasia. Exclusion criteria included antibiotic administration within the last 10 days, vaginal bleeding, vaginal douching, the use of intra-uterine contraceptive devices, and digital examinations with antimicrobial agents. Additionally, women undergoing laparoscopic hysterectomy were not eligible to participate. Protocols were approved by the University of Padua. Methods were carried out in accordance with the relevant guidelines and regulations. Written informed consent was obtained from all participating women on or before admission.

**Sample collection.** Cervical, vaginal, rectal, and oral Dacron swabs were obtained from each woman within 24 hours of the hysterectomy procedure. During a speculum examination, a swab was gently rotated around the external cervical os (i.e., the opening of the cervical canal into the vagina) for approximately 30 seconds and then removed while avoiding contact with vaginal structures. A second swab was then used to collect vaginal fluid from the posterior fornix, again allowing 30 seconds for saturation of the swab. Transabdominal hysterectomy was performed, with intraoperative antibiotic prophylaxis not being administered until the uterus had been removed. Upon removal of the uterus, it was handled by a research nurse under sterile conditions in the operating room. A sterile knife was used to separate the cervix from the uterine corpus. The uterus was opened with sterile scissors at 3 and 9 o'clock, taking precautions to avoid contamination of the endometrium. Once the uterus was opened, a Dacron swab was collected from the middle portion of the endometrium, an area unlikely to be contaminated during the procedure of opening the uterus, by gently rotating and allowing for saturation of the swab. An analysis was performed on a subset of subjects for whom swabs of both the mid-endometrium and the whole-length endometrium were collected (N = 9). There were no differences in 16S rRNA gene abundance, alpha diversity, or beta diversity between these samples (Supplementary Methods). Therefore, the data presented in this study are from mid-endometrium samples. All swabs were placed in cryovials and frozen at  $-80^{\circ}\text{C}$  until analysis.

**Extraction of DNA from samples.** Genomic DNA was extracted from swab samples using a QIAGEN DNeasy PowerLyzer PowerSoil Kit with minor modifications to the manufacturer's protocol. Within the supplied bead tube, swabs were immersed in 500  $\mu\text{l}$  of bead solution and 200  $\mu\text{l}$  of phenol:chloroform:isoamyl alcohol (pH 7–8) solution for 10 minutes. Sixty  $\mu\text{l}$  of Solution C1 were added, and microbial cells were lysed by mechanical disruption using a bead beater (twice at 30 seconds each). The bead tubes were centrifuged, and the supernatants were transferred to new tubes. Next, 100  $\mu\text{l}$  of solution C2, 100  $\mu\text{l}$  of solution C3, and one  $\mu\text{l}$  of RNase A enzyme were added and incubated at  $4^{\circ}\text{C}$  for five minutes. Steps involving solutions C2 and C3 were combined to maximize DNA yield. Tubes were centrifuged, and supernatants were transferred to new tubes that contained 650  $\mu\text{l}$  of solution C4 and 650  $\mu\text{l}$  of 100% ethanol. The lysates were loaded onto filter columns, centrifuged for one minute, and the flow-through was discarded. This step was repeated until all sample lysates were spun through the filter columns. Five hundred  $\mu\text{l}$  of solution C5 were added to the filter columns, centrifuged for one minute, the flow-through was discarded, and the tube was centrifuged for an additional two minutes as a dry-spin. Finally, 60  $\mu\text{l}$  of solution C6 were placed on the filter column and incubated for five minutes before centrifuging for 30 seconds to elute the extracted DNA. For each set of extractions, one blank DNA extraction kit was processed as a background negative control. Purified DNA was stored at  $-20^{\circ}\text{C}$ .



**Figure 1.** Quantitative real-time PCR (qPCR) analysis illustrating differences in 16S rRNA gene abundance based on cycle of quantification (Cq) values among oral, rectal, vaginal, cervical, endometrial, and technical control samples, including nuclease-free water. Bars indicate mean values.

**Quantitative real-time PCR (qPCR) of the 16S rRNA gene in samples.** Three qPCR assays targeting the 16S rRNA genes in samples were performed. Total bacterial DNA abundance within samples was measured via amplification of the V1 - V2 region of the 16S rRNA gene according to the protocol of Dickson *et al.*<sup>45</sup> with minor modifications. These modifications included the use of a degenerative forward primer (27f-CM: 5'-AGA GTT TGA TCM TGG CTC AG-3')<sup>46</sup> and a degenerate probe containing locked nucleic acids (+) (BSR65/17: 5'-56FAM-TAA +YA+C ATG +CA+A GT+C GA-BHQ1-3'). Each 20  $\mu$ l reaction contained 0.6  $\mu$ M of 27f-CM primer, 0.6  $\mu$ M of 357 R primer (5'-CTG CTG CCT YCC GTA G-3'), 0.25  $\mu$ M of BSR65/17 probe, 10.0  $\mu$ l of 2X TaqMan Environmental Master Mix 2.0 (Life Technologies, Carlsbad, CA), and 4.0  $\mu$ l purified DNA. The total bacterial DNA qPCR was performed under the following conditions: 95  $^{\circ}$ C for 10 min, followed by 40 cycles of 94  $^{\circ}$ C for 30 sec, 57  $^{\circ}$ C for 30 sec, and 72  $^{\circ}$ C for 30 sec.

Absolute abundances of *Lactobacillus iners* and *L. crispatus* were measured within a subset of vaginal, cervical, and endometrial samples from 10 and 8 women, respectively. Samples from these women were selected because the relative abundance of *L. iners* or *L. crispatus* in the bacterial profile of their vaginal sample was representative of the spectrum of relative abundances (~ 0–100%) of these bacteria observed among vaginal samples overall. Absolute abundances were measured via qPCR assays using taxon-specific primers described by Fredricks *et al.*<sup>47</sup> (*L. crispatus*: 5'-TCT TGA CAT CTA GTG CCA TTT GT-3' and 5'-TGC ACC ACC TGT CTT AGC-3') and Srinivasan *et al.*<sup>48</sup> (*L. iners*: 5'-GAT GCT AAT ACC GGA TAA YAA CAG AT-3' and 5'-CAC CGC AGG TCC ATC CAA GA-3'). For both assays, each 20  $\mu$ l reaction contained 1.0  $\mu$ M of each primer, 10.0  $\mu$ l of 2X Powerup SYBR Green Master Mix (Thermo Fisher Scientific, Carlsbad, CA), and 2.0  $\mu$ l purified DNA. Cycling conditions for the *L. iners* assay were: 95  $^{\circ}$ C for 10 min, and 45 cycles of 94  $^{\circ}$ C for 30 sec, 57  $^{\circ}$ C for 30 sec, and 72  $^{\circ}$ C for 30 sec. Cycling conditions for the *L. crispatus* assay were: 95  $^{\circ}$ C for 2 min, and 45 cycles of 95  $^{\circ}$ C for 15 sec and 60  $^{\circ}$ C for 1 min. For the taxon-specific qPCR assays, all samples were run in duplicate in a single run (total of two runs).

For all qPCR assays, raw amplification data were normalized to the ROX passive reference dye and analyzed with Standard Curve 3.3.0-SR2-build15 (Thermo Fisher Cloud), using automatic threshold and baseline settings. Cycle of quantification (Cq) values were calculated for samples based on the mean number of cycles required for normalized fluorescence to exponentially increase. In an effort to limit analyses to samples that produced bacterial signals beyond those evident in background technical controls, only samples of body sites that produced Cq values less than 35 for the total 16S rRNA gene qPCR (V1-V2) analysis were included in downstream 16S rRNA gene sequence-based analyses (Fig. 1).

To assess differences in 16S rDNA abundance between endometrial, cervical, vaginal, rectal, and oral samples among the 25 subjects, differences in qPCR Cq values were evaluated via repeated measures ANOVA followed by Tukey's pairwise comparisons<sup>49,50</sup>. *F*-tests were used to evaluate unequal variances in Cq values between body sites and background technical samples. To assess differences in 16S rDNA abundance between individual body sites and background technical control samples, *t*-tests or Mann-Whitney *U*-tests were used. Statistical analyses were performed using PAST software (v2.17c)<sup>51</sup>.

**Sequencing of the 16S rRNA gene in samples.** The V4 region of the 16S rRNA gene was selected for analysis because it is an established marker for broad-level phylogenetic analysis of bacterial taxa and it is the region of the 16S rRNA gene most representative of the nearly full-length gene<sup>52</sup>. Sequencing of the V4 region has been previously used to profile the bacterial communities of multiple human body sites, including the skin<sup>53</sup>, mouth<sup>54</sup>, lung<sup>55</sup>, gut<sup>56</sup>, vagina<sup>57</sup>, bladder<sup>58</sup>, placenta<sup>59</sup>, and endometrium<sup>25,33</sup>. Amplification and sequencing of the V4 region of the 16S rRNA gene was performed at the University of Michigan's Center for Microbial Systems (Ann Arbor, MI) using the dual indexing sequencing strategy developed by Kozich *et al.*<sup>60</sup>. Sequencing was conducted using the Illumina MiSeq platform (V2 500 cycles, Illumina MS102-2003), according to the manufacturer's instructions with modifications found in Kozich *et al.*<sup>60</sup>. Each PCR reaction contained 1.0  $\mu$ M of each primer, 5.0  $\mu$ l template DNA, 0.15  $\mu$ l AccuPrime HiFi Polymerase, and DNase-free water to produce a final volume of 20  $\mu$ l. Standard PCR was performed using the following conditions: 95 °C for 2 minutes, followed by 30 cycles of 95 °C for 20 seconds, 55 °C for 30 seconds, and 72 °C for 5 minutes, with an additional elongation at 72 °C for 10 minutes.

The resulting fragments were visualized in 3% agarose gel. Several cervical, endometrial, and blank DNA extractions kit samples produced no or weak bands, indicating that sequencing of these samples was unlikely to yield sufficient sequence data for downstream analysis. Therefore, DNA amplification was additionally performed on cervical, endometrial, and technical control samples using touchdown PCR. This method reduces the initial amplification of nonspecific host DNA sequences during early steps of amplification by using a relatively high primer annealing temperature in relation to the melting point of the primers, and incrementally decreasing the annealing temperature as cycling proceeds<sup>61</sup>. This method has been recently used to investigate the microbiota of several body sites with low microbial biomass, including the lung<sup>62,63</sup> and brain<sup>64</sup>.

Touchdown PCR was performed using the following conditions: 2 min at 95 °C, followed by 20 cycles of 95 °C for 20 s, 60 °C for 15 s, and 72 °C for 5 min (with a 0.3 °C decrease of the 60 °C annealing temperature each cycle), followed by 20 cycles of 95 °C for 20 s, 55 °C for 15 s, and 72 °C for 5 min, followed by 72 °C for 10 min. Three  $\mu$ l of template DNA were included in each touchdown PCR reaction. Sequencing libraries were prepared according to Illumina's protocol for Preparing Libraries for Sequencing on the MiSeq (15039740 Rev. D) for 2 nM or 4 nM libraries. FASTQ files were generated for paired-end reads. Sample-specific MiSeq run files have been deposited on the NCBI Sequence Read Archive (BioProject ID PRJNA543861).

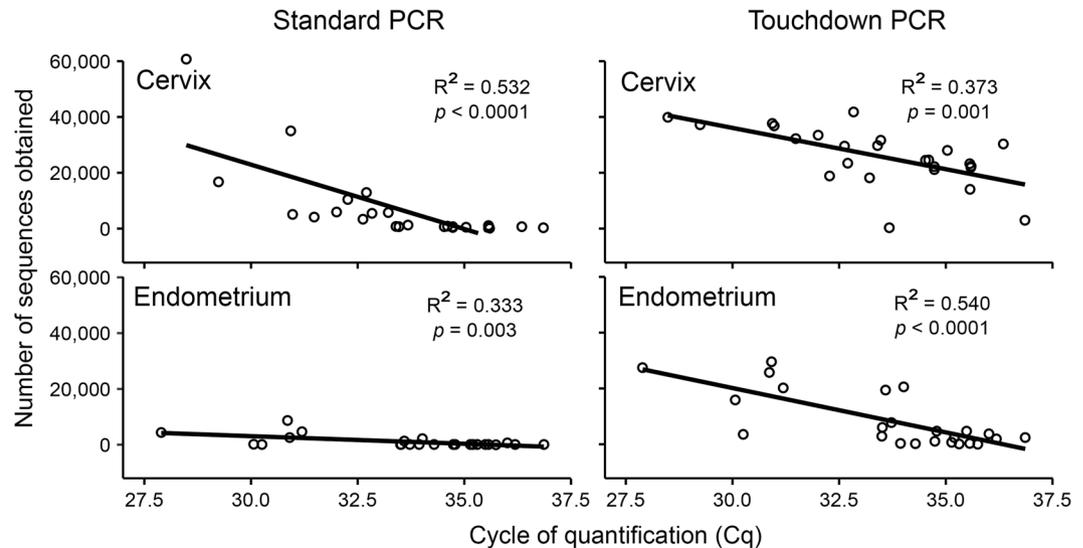
**16S rRNA gene sequence processing and data analysis.** Prior to analyzing bacterial profile alpha and beta diversity, raw sequenced reads were processed using Mothur software (v1.39.5)<sup>65</sup>, following the Standard Operating Procedure provided by Schloss *et al.* ([www.mothur.org/wiki/MiSeq\\_SOP](http://www.mothur.org/wiki/MiSeq_SOP)). Paired-end reads were assembled into contiguous sequences, quality checked (maximum length = 300, maximum ambiguous base pairs = 0, and maximum number of homopolymers = 8), and aligned against the SILVA 16S rDNA reference database (release 102); sequences falling outside the target alignment space were removed<sup>66,67</sup>. Quality sequences were pre-clustered (diffs = 2) and chimeric sequences were identified with VSEARCH and removed<sup>68</sup>. The remaining sequences were taxonomically classified using the SILVA reference database<sup>67</sup> with a k-nearest neighbor approach and a confidence threshold of 80%. Sequences derived from an unknown domain, Eukaryota, Archaea, chloroplasts, or mitochondria were removed. Operational taxonomic units (OTUs) were defined by clustering sequences at a 97% sequence similarity cutoff using the average neighbor method. Sequencing of mid-endometrial, whole-length endometrial, cervical, vaginal, rectal, oral, and background technical control samples yielded 5,715,294 sequences. They clustered into 3,338 OTUs (1,019 singletons, and 517 doubletons). The median number of sequences per sample was 16,676 (13,026–20,906 95% CI). The raw OTU data from this study are provided as Supplementary Data.

Heatmaps were generated for prominent OTUs (i.e., average relative abundance  $\geq 1\%$ ) using the open-source software program Morpheus (<https://software.broadinstitute.org/morpheus>). For determining the taxonomic identities of select OTUs beyond the genus level, resultant consensus sequences were submitted for Basic Local Alignment Search Tool (BLAST)<sup>69</sup> analysis. To investigate linear relationships between absolute 16S rDNA abundances and the relative abundances of individual bacterial taxa in sequencing surveys, Spearman's correlation tests were used.

Alpha diversity analyses were based on three metrics: Chao 1 richness estimator, Shannon diversity index, and the inverse Simpson index, each calculated using Mothur software (v1.39.5)<sup>65</sup>. Differences in alpha diversity values between body sites and background technical controls samples were evaluated using Mann-Whitney *U*-tests in R (v3.5.1)<sup>70</sup>. Differences in alpha diversity values among paired body site samples were evaluated using linear mixed-effect models and ANOVA tests, controlling for subject (i.e. patient) identity as a random effect using the R-package lme4<sup>71</sup>, followed by Tukey's pairwise comparisons with Bonferroni adjustments using the R-package multcomp<sup>72</sup>.

Beta diversity was assessed using Jaccard and Bray-Curtis similarity indices to reflect bacterial profile composition and structure, respectively. Similarity values were calculated using percent relative abundance data for OTUs within samples. Beta diversity was visualized through Principal Coordinates Analyses (PCoA), and statistically evaluated using PERMANOVA with 9,999 permutations<sup>73</sup>. The influences of subject identity and body site in structuring the microbiota were concurrently investigated using the *adonis* function in the R-package vegan (v2.4-6)<sup>74</sup>. Linear discriminant analysis effect size, or LEfSe<sup>75</sup>, was used with default parameters ( $\alpha = 0.05$  and LDA score 2.0) to identify any OTUs that differed in relative abundance between body sites and/or background technical controls. Prior to LEfSe analysis, singleton and doubleton OTUs were removed from the dataset.

Analysis of the bacterial profiles of body site samples was performed only on samples with a qPCR C<sub>q</sub> value less than 35 (all blank DNA extraction kit controls had a C<sub>q</sub> value  $\geq 34.9$ ). Additionally, analyses were limited to body site samples that had Good's coverage values greater than 98% and contained at least 500 (standard PCR



**Figure 2.** Regressions of the number of quality 16S rRNA gene bacterial sequences obtained from the cervix and endometrium using standard PCR and touchdown PCR approaches against the absolute abundance of 16S rDNA in samples based on cycle of quantification (Cq) values from universal bacterial quantitative real-time PCR.

dataset) or 1,000 (touchdown PCR dataset) sequences. For analyses of alpha diversity, sequence libraries were subsampled to a depth of either 500 or 1,000 sequences, for the standard PCR and touchdown PCR data sets, respectively. For beta diversity analyses, no subsampling was performed. Background technical controls were included in beta diversity analyses only if they had Good's coverage values greater than 90% and yielded at least 100 quality sequences.

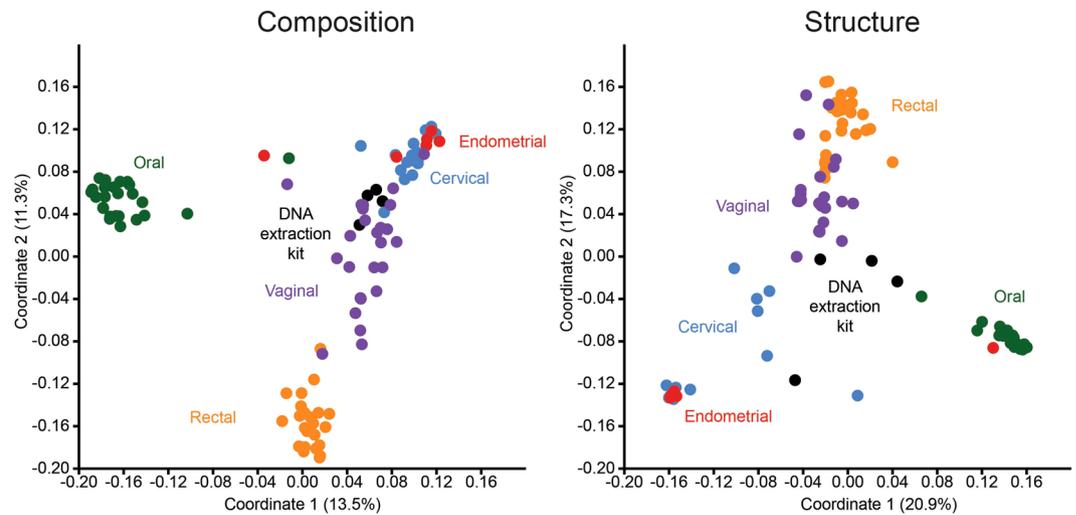
A *post-hoc* analysis was conducted on the 16S rRNA gene sequences from cervical, endometrial, and technical control samples using the Dada2 (version 1.8) package<sup>76</sup> in R (v3.5.1)<sup>70</sup>, and the online MiSeq protocol (<https://benjjneb.github.io/dada2/tutorial.html>). This secondary analysis of amplicon sequence variants (ASVs), defined by 100% sequence similarity, may afford higher resolution of prominent sequence variants among body site and technical control samples. Amplicon sequence variant analysis was only performed on cervical and endometrial samples with a qPCR Cq value less than 35, and which had at least 200 sequences with a Good's coverage value > 90% for the standard PCR dataset, or at least 420 sequences and a Good's coverage value > 97% for the touchdown PCR dataset. Amplicon sequence variant analysis was performed on all background DNA control samples.

## Results

**Absolute 16S rDNA abundance (i.e. bacterial load) among body sites.** Quantitative real-time PCR revealed variation in 16S rDNA abundance across body sites (Fig. 1). Oral, rectal, and vaginal samples had lower cycle of quantification (Cq) values (i.e. higher bacterial loads) than cervical and endometrial samples (repeated measures ANOVA:  $df = 124$ ,  $F = 313.7$ ,  $p < 0.0001$ ; Tukey's pairwise comparisons:  $p = 0.0001$ ). The median fold change in Cq value in vaginal samples compared to cervical and endometrial samples was 1.75 (1.60–1.89 95% CI) and 1.77 (1.62–1.90 95% CI), respectively. Cervical and endometrial samples did not differ in Cq value ( $t = 0.69$ ;  $p = 0.49$ ), but each had a lower Cq value than background technical controls ( $p \leq 0.02$ ) (Fig. 1). Nevertheless, the Cq values of cervical and endometrial samples were variable, with only some (33/50) of these samples exhibiting bacterial molecular signals exceeding those of technical controls (Fig. 1). Specifically, the range of Cq values for blank DNA extraction kit controls was 34.9–36.6 cycles; 72% (18/25) of cervical and 60% (15/25) of endometrial samples had Cq values lower than those of extraction controls. Bacterial load in the vagina was not correlated with bacterial load in either the cervix ( $R^2 = 0.02$ ,  $p = 0.46$ ) or the endometrium ( $R^2 < 0.0001$ ,  $p = 0.99$ ). However, bacterial load in the cervix was positively correlated with that in the endometrium ( $R^2 = 0.32$ ,  $p = 0.003$ ).

For the cervix and endometrium, significant correlations were observed between 16S rDNA abundances in the samples, as determined by qPCR, and in the number of quality trimmed bacterial sequences obtained from MiSeq sequencing using both standard PCR and touchdown PCR approaches for generating sequence libraries (Fig. 2). These results demonstrate that the number of quality sequences obtained was directly related to the absolute abundances of 16S rDNA in samples, providing confidence that the sequence data from cervical and endometrial samples were not simply due to background DNA contamination (i.e. the “kitome”).

**Bacterial profiles among body sites (standard PCR dataset).** In general, rectal and oral samples had higher levels of taxonomic richness (Chao1 estimator) and heterogeneity (Shannon and Inverse Simpson indices) than vaginal, cervical, and endometrial samples (Supplementary Fig. S1, Table S2). Specifically, the richness and heterogeneity of rectal samples exceeded those of all other sample types. Additionally, the heterogeneity of oral



**Figure 3.** Principal Coordinates Analyses (PCoA) illustrating differences in 16S rRNA gene profiles among oral, rectal, vaginal, cervical, endometrial, and DNA extraction kit samples using a standard PCR approach. Profiles were generated for 16S rRNA gene community composition and structure using the Jaccard and Bray-Curtis indices, respectively.

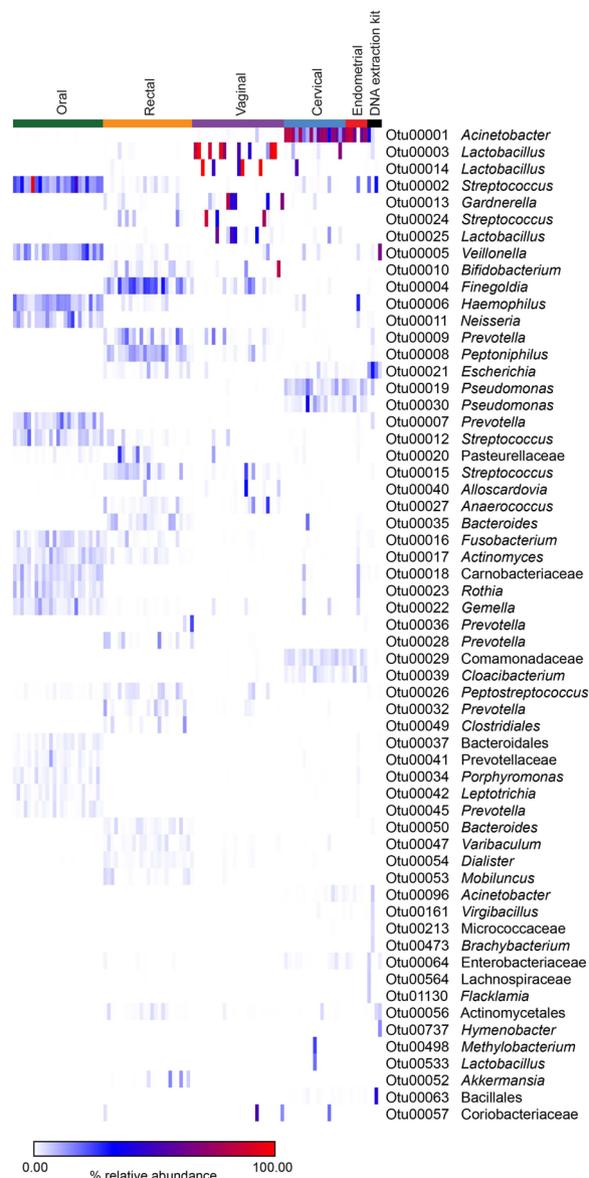
samples was greater than that of vaginal, cervical, and endometrial samples, and the richness of oral samples was greater than that of vaginal and cervical samples, but not that of endometrial samples. Lastly, the richness and heterogeneity of vaginal, cervical, and endometrial samples did not vary, with the exception of cervical samples having higher Shannon diversity indices than vaginal samples (Supplementary Fig. S1, Table S2).

With respect to beta diversity, there was a high degree of separation among the bacterial profiles of body sites based on community composition (Jaccard Similarity Index;  $F = 9.49$ ,  $p = 0.0001$ ) and structure (Bray-Curtis Similarity Index;  $F = 19.20$ ,  $p = 0.0001$ ) (Fig. 3). Pairwise analyses with sequential Bonferroni corrections showed that, with the exception of the cervix and endometrium (whose profiles did not consistently differ), the bacterial profiles of the body sites differed in both composition and structure ( $p = 0.0001$ ). Although the bacterial profiles of the cervix and endometrium did not differ from one another, the profiles of each differed from those of background technical controls in both composition ( $p \leq 0.005$ ) and structure ( $p \leq 0.02$ ) (Fig. 3). When only subjects with paired vaginal, cervical, and endometrial samples that met the cutoff criteria (500 sequences and Good's coverage  $\geq 98\%$ ) were considered ( $N = 6$ ), and after controlling for subject (i.e. patient) identity, the effect of body site remained for community composition (Jaccard: subject,  $R^2 = 0.301$ ,  $p = 0.017$ ; sample type,  $R^2 = 0.127$ ,  $p = 0.005$ ), and for community structure (Bray-Curtis: subject,  $R^2 = 0.293$ ,  $p = 0.085$ ; sample type,  $R^2 = 0.275$ ,  $p = 0.001$ ).

Investigation of the taxonomic identities of prominent OTUs (defined as OTUs with  $\geq 1\%$  average relative abundance) by sample type revealed overlap in the profiles of cervical and endometrial samples compared with those of vaginal, rectal, and oral samples (Fig. 4). The profiles of vaginal samples were dominated by *Lactobacillus* and *Gardnerella*. The most prominent taxa in rectal samples were *Finegoldia*, *Prevotella*, *Peptoniphilus*, *Streptococcus*, and *Bacteroides*. In oral samples, the most prominent taxa were *Streptococcus*, *Veillonella*, *Haemophilus*, *Neisseria*, *Prevotella*, *Fusobacterium*, *Actinomyces*, *Rothia*, and *Gemella*. In contrast, the bacterial profiles of both cervical and endometrial samples were dominated by *Acinetobacter*, with a single prominent OTU (OTU 1) having average relative abundances of 53.3% in the cervix and 60.6% in the endometrium. Core OTUs (defined as OTUs present in  $\geq 50\%$  of samples, and with a mean relative abundance  $\geq 1\%$ ) in both the cervix and endometrium included *Acinetobacter* (OTU 1), *Pseudomonas* (OTUs 19 and 30), Comamonadaceae (OTU 29), and *Cloacibacterium* (OTU 39). Notably, *Lactobacillus* (OTU 3) was a core OTU in the cervix but not in the endometrium. The two most abundant *Lactobacillus* OTUs in the dataset (OTUs 3 and 14) accounted for 26,462/137,551 (19.24%) sequences among cervical samples, and yet only 4/23,754 (0.017%) sequences among endometrial samples.

A BLAST analysis (i.e. comparison of a 16S rRNA gene sequence to those in the BLAST taxonomy database) revealed that the consensus sequence for OTU 1 was identical to that of eight *Acinetobacter* type strains within seven species that have been isolated from clinical samples: *A. bereziniae*, *A. colistiniresistens*, *A. gyllenbergii*, *A. junii*, *A. modestus*, *A. proteolyticus*, and *A. vivianii*. A BLAST analysis of OTU 30 showed identical matches for multiple strains of *Pseudomonas aeruginosa*, and an analysis of OTU 19 revealed identical matches for multiple *Pseudomonas* species within the *Pseudomonas putida* group (*P. monteilii*, *P. plecoglossicida*, *P. putida*, and *P. taiwanensis*)<sup>77</sup>. A BLAST analysis of OTU 39 showed identical matches for six type strains within the species *Cloacibacterium normanense*, *C. caeni*, and *C. rupense*. Lastly, a BLAST analysis of OTU 29 (a member of the family Comamonadaceae) revealed identical matches for *Comamonas jiangduensis* and *C. kerstersii*.

The most prominent taxa in background technical controls were *Veillonella* (OTU 5), *Escherichia* (OTU 21), *Streptococcus* (OTU 2), and *Acinetobacter* (OTU 1) (Fig. 4). However, *Escherichia* (OTU 21) was the only genus consistently detected in high relative abundances across the technical control samples (12.4% to 31.3%, in three of four controls).

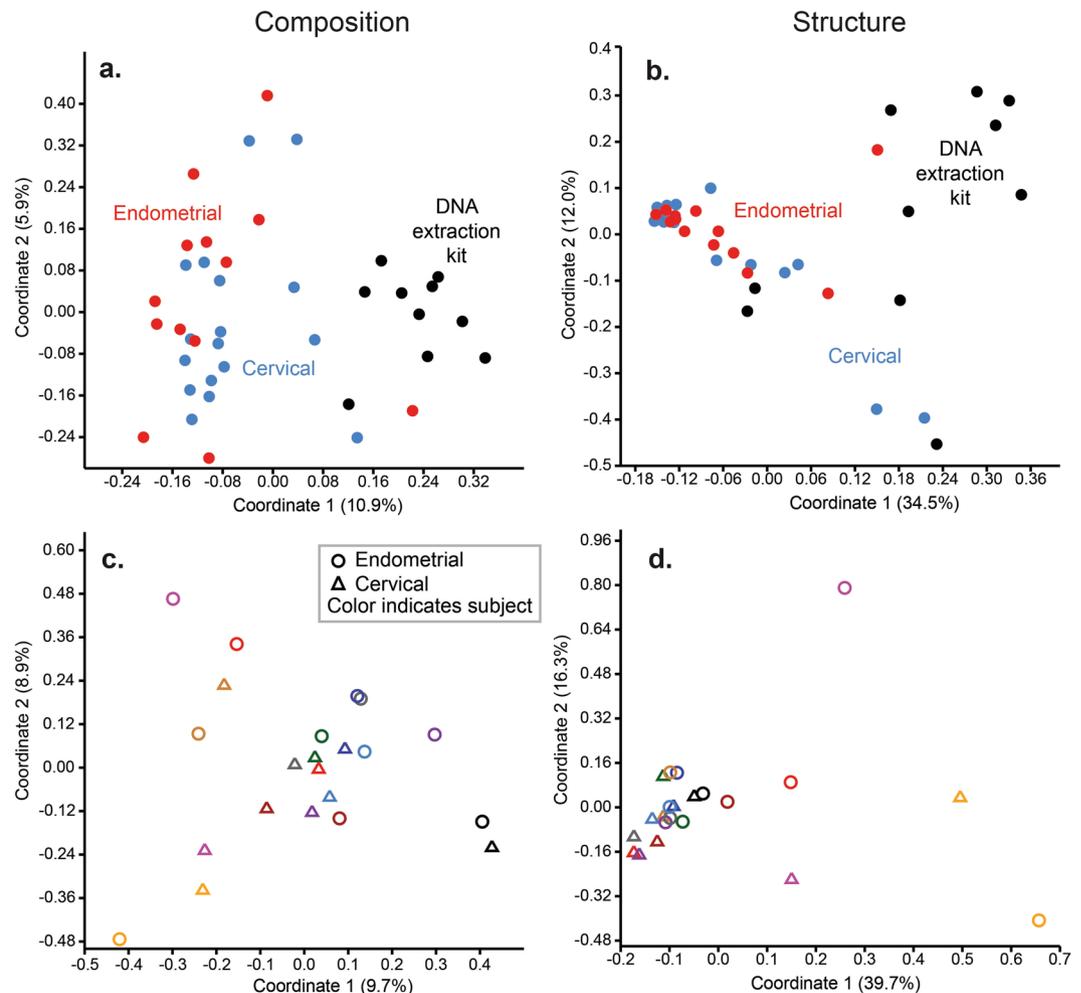


**Figure 4.** Heat map illustrating percent relative abundances of prominent operational taxonomic units ( $\geq 1\%$  average relative abundance) among oral, rectal, vaginal, cervical, endometrial, and DNA extraction kit samples. Amplification of 16S rRNA genes was performed using a standard PCR approach.

**Bacterial profiles of the cervix and endometrium (touchdown PCR dataset).** For the touchdown PCR dataset, analysis of alpha diversity showed that the richness of the bacterial profiles of the cervix and endometrium did not differ from those of background technical controls (Supplementary Fig. S1, Table S3). However, both the cervix and endometrium had greater bacterial profile heterogeneity than the technical controls (Supplementary Table S3).

With respect to beta diversity, there was separation between the composition and structure of the bacterial profiles of cervical and endometrial samples and those of technical controls (Fig. 5a,b; PERMANOVA,  $p = 0.0001$ ). However, the bacterial profiles of the cervix and endometrium did not consistently differ from each other ( $p > 0.05$ ). Instead, subject (i.e. patient) identity explained most of the variation in bacterial profiles among the cervical and endometrial samples (Jaccard: subject,  $R^2 = 0.49$ ,  $p = 0.002$ , sample type  $R^2 = 0.04$ ,  $p = 0.50$ ; Bray-Curtis: subject,  $R^2 = 0.56$ ,  $p = 0.001$ , sample type,  $R^2 = 0.05$ ,  $p = 0.09$ ) (Fig. 5c,d).

Investigation of the taxonomic identities of prominent OTUs ( $\geq 1\%$  average relative abundance) by sample type in the touchdown PCR dataset (Fig. 6) was congruent with analyses for the standard PCR dataset. Specifically, the bacterial profiles of cervical and endometrial samples were dominated by *Acinetobacter* (OTU 1), which accounted for 49.0% and 44.4% of sequences from these body sites, respectively. Other prominent taxa in the endometrium included *Pseudomonas* (OTUs 19 and 30), *Cloacibacterium* (OTU 39), Comamonadaceae (OTU 29), and *Escherichia* (OTU 21). Other prominent taxa in the cervix included *Pseudomonas* (OTUs 19 and 30), *Cloacibacterium* (OTU 39), Comamonadaceae (OTU 29), *Lactobacillus* (OTUs 3 and 14), *Escherichia* (OTU



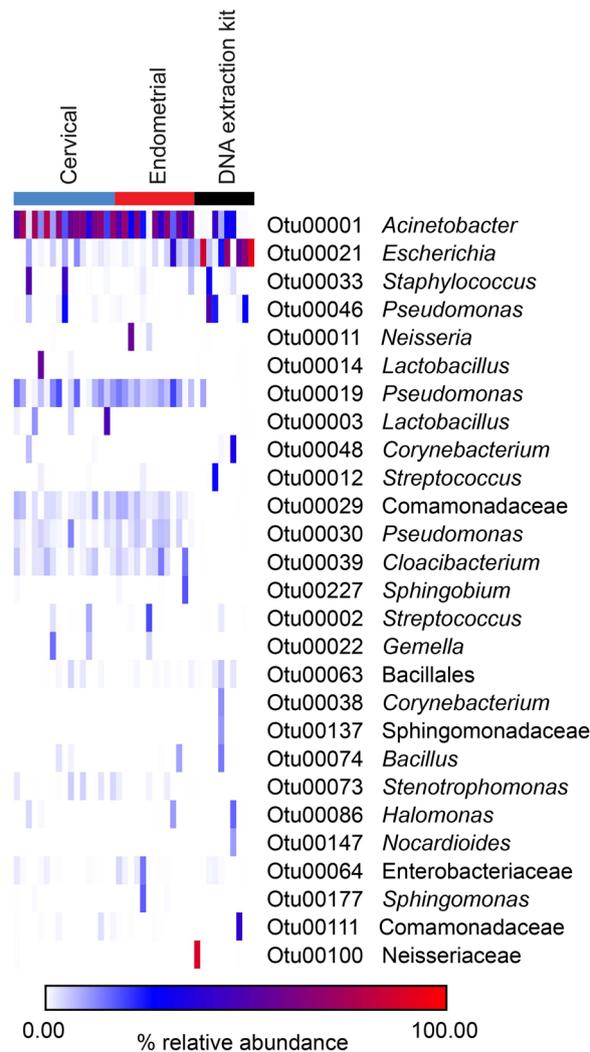
**Figure 5.** Principal Coordinates Analyses (PCoA) illustrating differences in 16S rRNA gene profiles among cervical, endometrial, and DNA extraction kit samples using a touchdown PCR approach. Profiles were generated for 16S rRNA gene community composition and structure using the Jaccard and Bray-Curtis indices, respectively. In panels c and d, the color coding of circles (endometrial) and triangles (cervical) indicates subject (i.e. patient) identity.

21), and *Staphylococcus* (OTU 33). Nine different *Lactobacillus* OTUs were identified in 14/17 (82.4%) cervical samples (one to three *Lactobacillus* OTUs per sample), wherein they accounted for 42,346/436,436 (9.7%) sequences. Among endometrial samples, a maximum of one *Lactobacillus* OTU (either OTU 3, OTU 14, or OTU 25) was detected in 3/13 (23.1%) samples. These three *Lactobacillus* OTUs accounted for only 11/62,325 (0.018%) sequences among these three endometrial samples.

The most widespread and abundant genus in the profiles of background technical controls was *Escherichia* (OTU 21), with an average relative abundance of 41.3% (Fig. 6). Other prominent taxa in the control samples were *Acinetobacter* (OTU1; 11.5%), *Pseudomonas* (OTU 46; 10.4%), and *Staphylococcus* (OTU 33; 3.6%).

Linear discriminant analysis effect size (LEfSe) identified OTUs that were differentially abundant between the cervix, endometrium, and background technical controls (Fig. 7 and Supplementary Table S4). Twenty-seven OTUs had differential relative abundances between the cervical and control samples, and 31 OTUs had differential relative abundances between endometrial and control samples. *Acinetobacter* (OTU 1), *Pseudomonas* (OTUs 19 and 30), Comamonadaceae (OTU 29), and *Cloacibacterium* (OTU 39) were most differentially abundant between the cervical and endometrial samples and background technical controls. Four OTUs had differential relative abundances between cervical and endometrial samples: *Lactobacillus*, *Stenotrophomonas*, and *Schlegelella* (OTUs 3, 219, and 134, respectively) were more relatively abundant in the cervix than the endometrium, and *Sphingobacterium* (OTU 110) was more abundant in the endometrium than the cervix.

**Amplicon sequence variant analysis of the cervix, endometrium, and background technical controls.** Amplicon sequence variants within 0–1 base pair similarity of the consensus sequences for *Acinetobacter* (OTU 1), *Pseudomonas* (OTUs 19 and 30), Comamonadaceae (OTU 29), and *Cloacibacterium* (OTU 39) were identified in cervical, endometrial, and background technical control samples (Supplementary Fig. S2). These five OTUs were prominent among cervical and endometrial samples (Fig. 6), and were indicated by LEfSe analysis as being relatively more abundant in both cervical and endometrial samples than in background

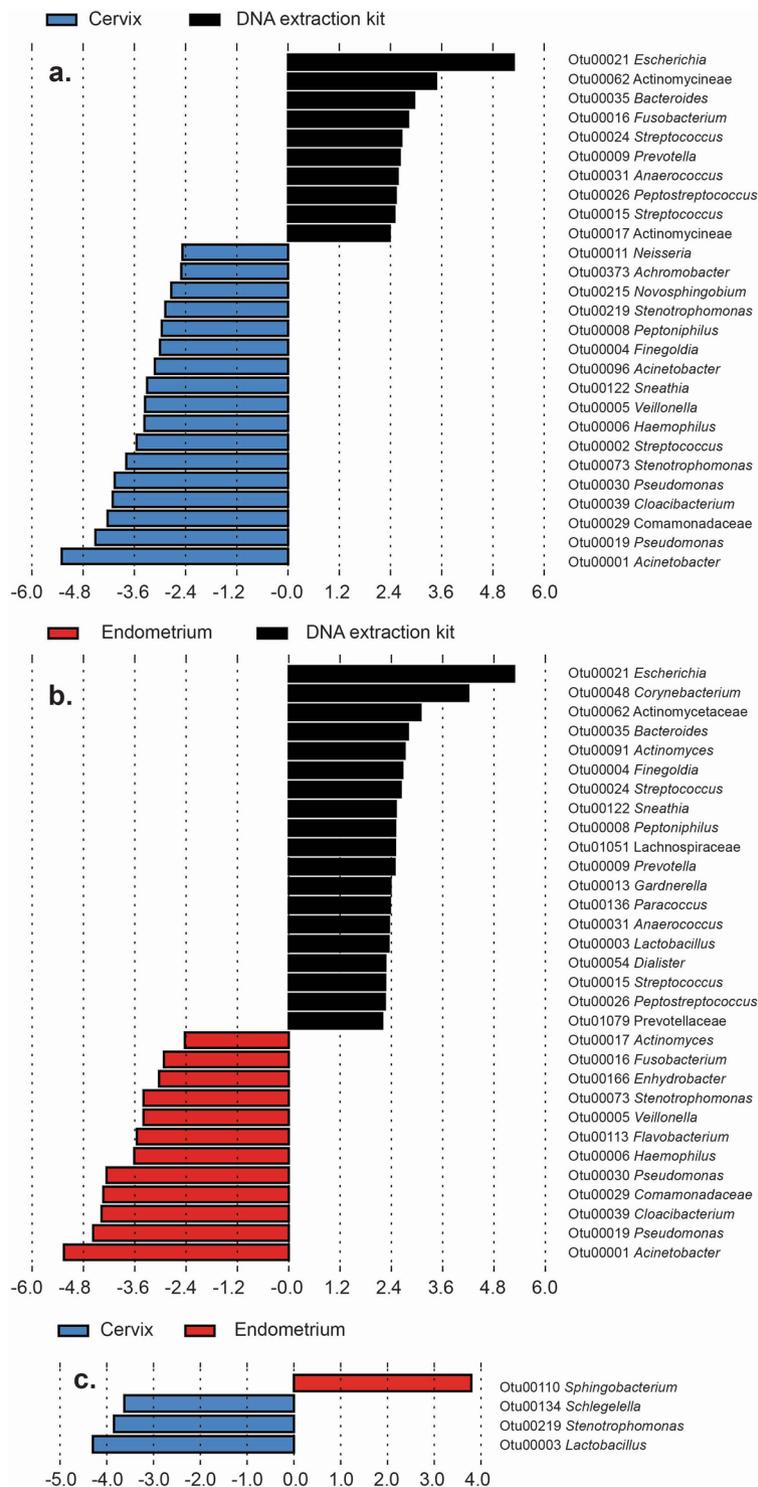


**Figure 6.** Heat map illustrating percent relative abundances of prominent operational taxonomic units ( $\geq 1\%$  average relative abundance) among cervical, endometrial, and DNA extraction kit samples. Amplification of 16S rRNA genes was performed using a touchdown PCR approach.

technical controls (Fig. 7). In the standard PCR dataset, amplicon sequence variants differing at most by one base pair from the consensus sequences of these prominent OTUs were widespread and relatively abundant among cervical and endometrial samples, yet they were rarely identified in technical controls (Supplementary Fig. S2). In the touchdown PCR dataset, amplicon sequence variants differing at most by one base pair from the consensus sequences of these prominent OTUs were again either widespread and relatively abundant among cervical and endometrial samples, yet rarely identified in technical controls (OTUs 30, 29, and 39), or they were widespread among cervical, endometrial, and technical controls, yet present at much higher relative abundances among the cervical and endometrial samples (OTUs 1 and 19) (Supplementary Fig. S2). Most notably, the amplicon sequence variant that was identical to the consensus sequence of OTU 1 was detected in all cervical samples and in 14/15 endometrial samples, with average relative abundances of 44.9% and 32.7%, respectively (Supplementary Fig. S2). This sequence variant was detected in each of the eight control samples, but at much lower relative abundances (0.03% to 1.15%) than in the cervical and endometrial samples.

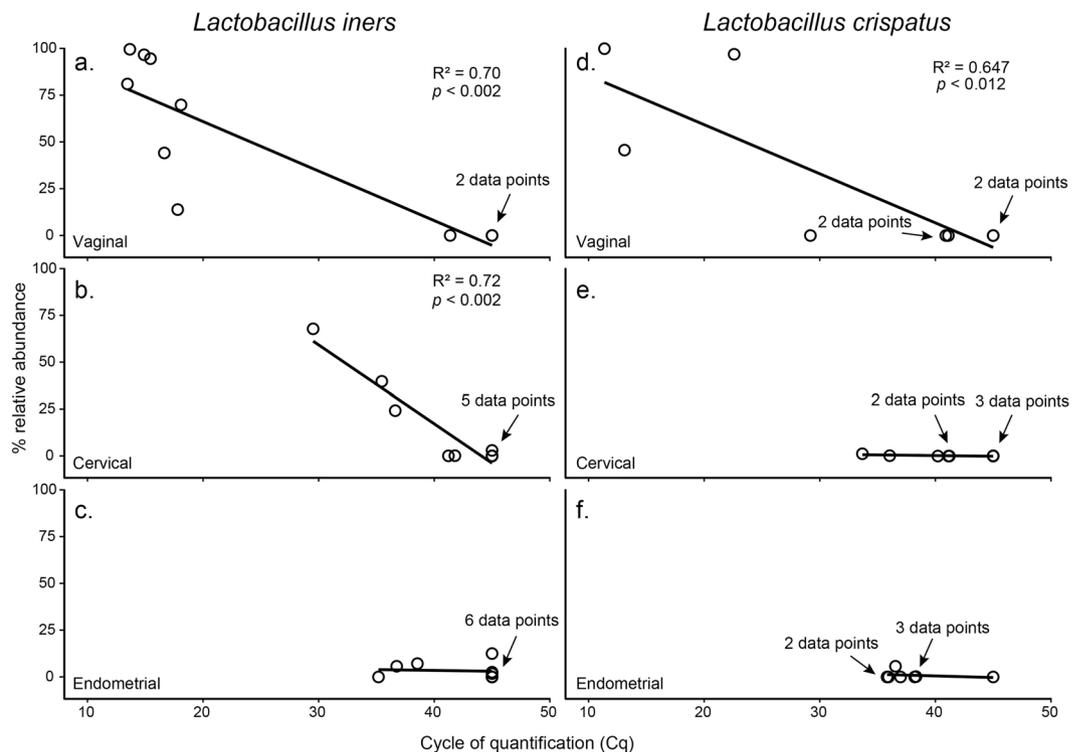
***Lactobacillus*-specific quantitative real-time PCR of vaginal, cervical, and endometrial samples.** Two targeted qPCR assays were performed to investigate the presence and absolute abundance of *L. iners* and *L. crispatus* in cervical and endometrial samples. Representative subjects (i.e. patients) were selected for analysis based on the relative abundances of OTU 3 (*L. iners*) and OTU 14 (*L. crispatus*) in their vaginal samples. Specifically, representative samples were selected to cover a broad range of relative abundances within the vagina for *L. iners* (N = 10; 0.02–99.5%) and *L. crispatus* (N = 8; 0.001–99.7%). Each subject's vaginal, cervical, and endometrial samples were assayed for the absolute abundance of the respective *Lactobacillus* species. The qPCR assays demonstrated an absence or very low abundance of *Lactobacillus* in the endometrium, even when *Lactobacillus* was abundant in the vagina (Fig. 8).

For *L. iners*, a strong correlation between the relative (MiSeq sequence data) and absolute (qPCR Cq data) abundances was observed for vaginal (Fig. 8a) and cervical samples (Fig. 8b), but not for endometrial samples



**Figure 7.** Linear discriminant analysis effect size (LEfSe) of bacterial operational taxonomic units (OTUs) with significant differential abundance in cervical, endometrial, and DNA extraction kit samples. Classification of differential OTUs between (a) the cervix and DNA extraction kits, (b) the endometrium and DNA extraction kits, and (c) the cervix and endometrium.

(Fig. 8c). The *L. iners* assay resulted in positive amplification ( $C_q < 35$ ) for 7/10 (70%) vaginal samples, 1/10 (10%) cervical samples, and 0/10 (0%) endometrial samples. The vaginal sample with the highest relative abundance for *L. iners* (99.5%) in its respective MiSeq library had a  $C_q$  value of 13.7 for the *L. iners* qPCR assay. The three vaginal samples with the lowest relative abundances for *L. iners* (each < 1%) in their respective MiSeq libraries had  $C_q$  values > 35.



**Figure 8.** Regressions of the relative (MiSeq sequencing data) and absolute (qPCR Cq data) abundances of *Lactobacillus iners* and *L. crispatus* in vaginal, cervical, and endometrial samples.

For *L. crispatus*, a strong correlation between the relative (MiSeq) and absolute (qPCR) abundances was observed only in vaginal samples (Fig. 8d). The *L. crispatus* assay resulted in positive amplification for 4/8 (50%) vaginal samples, 1/8 (12.5%) cervical samples (Figs. 8e), and 0/8 (0%) endometrial samples (Fig. 8f). The vaginal sample with the highest relative abundance for *L. crispatus* (99.7%) in its respective MiSeq library had a Cq value of 11.4 for the *L. crispatus* qPCR assay. The three vaginal samples with the lowest relative abundances for *L. crispatus* (each <0.005%) in their respective MiSeq libraries had Cq values > 35. Thus, neither *L. iners* nor *L. crispatus* was widespread or abundant among endometrial samples (Fig. 8c,f).

## Discussion

The principal findings of the study were: 1) 60% (15/25) of the middle endometrial samples had a bacterial load exceeding that of background DNA controls; 2) Bacterial loads in the endometrium were correlated with those in the cervix ( $R^2 = 0.32$ ,  $p = 0.003$ ), but not with those in the vagina; 3) Among women with bacterial loads in the endometrium exceeding those of background controls, endometrial bacterial profiles were distinct from those of the controls and other body sites, except the cervix; 4) Endometrial and cervical bacterial profiles were similar when characterized using standard and touchdown PCR approaches: in both instances, they were dominated by *Acinetobacter*, *Pseudomonas*, *Cloacibacterium*, and Comamonadaceae; 5) *Lactobacillus* was dominant in vaginal bacterial profiles, and it was widespread among cervical profiles, yet it was rare in the bacterial profiles of endometrial samples using 16S rRNA gene sequencing (i.e. in 23% (3/13) of endometrial samples, accounting for 0.006% of sequence reads in these three samples, within the touchdown PCR dataset); and 6) Using species-specific qPCR, *Lactobacillus iners* and *L. crispatus* were detected (i.e.  $Cq < 35$ ) among vaginal samples with a spectrum of *L. iners* and *L. crispatus* relative abundances (13.8% – 99.5% and 0.008% – 99.7%, respectively), yet neither species was detected by qPCR in endometrial samples.

Recent studies concluding the existence of a resident endometrial microbiota through molecular techniques have reported varying taxonomic structures for this microbiota<sup>24–37,78</sup> (Supplementary Table S1). Most studies in which endometrial samples were collected transcervically have reported that *Lactobacillus*, a genus commonly dominant in the human vagina<sup>79–85</sup>, is also dominant in the endometrium<sup>25–28,33–37,78</sup>. The exception was a study by Verstraelen *et al.*<sup>29</sup>, which reported that 90% of women had endometrial bacterial profiles in which three *Bacteroides* and one *Pelomonas* species were most abundant. The presence and relative abundance of *Lactobacillus* in studies in which endometrial samples were not collected transcervically has been more variable<sup>24,30–32</sup>.

Using multiple species-specific qPCR assays to detect prevalent vaginal bacteria in vaginal, upper endocervical, and endometrial samples from 58 women who underwent hysterectomy, Mitchell *et al.*<sup>24</sup> determined that 95% of the women exhibited colonization of the endocervix and/or the endometrium by at least one bacterial species. The most commonly detected bacteria in the vagina were *Prevotella* spp. (76%), *L. iners* (61%), and *L. crispatus* (56%). *L. iners* (45%), *L. crispatus* (33%), and *Prevotella* spp. (33%) were also the most often detected bacteria among the endocervical/endometrial samples (endocervical and endometrial samples were considered

collectively as being representative of the upper genital tract)<sup>24</sup>. Chen *et al.*<sup>31</sup> recently used 16S rRNA gene sequencing to analyze the bacterial profiles of the lower and upper reproductive tracts of 110 women who underwent hysterectomy, and reported that the relative abundances of *Lactobacillus* decreased from the lower reproductive tract (median abundance of 99.9% in the vagina) to the upper reproductive tract (median abundance of 1.7% in the fallopian tubes). In the endometrium specifically, *Lactobacillus* remained relatively abundant (30.6%), but high relative abundances of *Acinetobacter* (9.1%), *Pseudomonas* (9.1%), *Vagococcus* (7.3%), *Sphingobium* (5.0%), and Comamonadaceae (4.9%) were also observed. Miles *et al.*<sup>32</sup> used 16S sequencing to investigate the presence of a microbiota in the reproductive tract of 10 women who underwent a total hysterectomy with bilateral salpingo-oophorectomy, and found that *Lactobacillus* was present in approximately half of the vaginal, cervical, endometrial, and myometrial samples at varying abundances. *Acinetobacter* and *Corynebacterium*, while not present in vaginal samples, were detected among cervical, endometrial, and myometrial samples. Lastly, Walther-Antônio *et al.*<sup>30</sup> used 16S sequencing to characterize the endometrial bacterial profiles of 10 women who underwent hysterectomy for non-malignant uterine conditions, and reported that *Lactobacillus* was infrequently detected, and then at only low relative abundances, within the endometrium. The most prominent bacteria in the endometrium were instead *Shigella* and *Barnesiella*. Therefore, although *Lactobacillus* has been reported as being widespread and dominant in endometrial samples collected transcervically, its distribution and abundance among endometrial samples that were not collected transcervically has been more variable.

In the current study, samples were obtained from the mid-endometrium after hysterectomy, and the endometrial microbiota was not dominated by *Lactobacillus*. Instead, endometrial bacterial profiles were largely dominated by *Acinetobacter*, *Pseudomonas*, Comamonadaceae, and *Cloacibacterium*. Each of these taxa has previously been detected in the human endometrium using molecular techniques<sup>25,26,28–33,35,36,78</sup> (Supplementary Table S4). In general, our findings are consistent with those of Chen *et al.*<sup>31</sup> and Miles *et al.*<sup>32</sup> in that we found decreasing levels of *Lactobacillus* from the vagina and the cervix to the endometrium. The rarity of *Lactobacillus* in the mid-endometrium in the current study suggests that the vaginal microbiota may not be a persistent source of microbes for the endometrial microbiota. This is further supported by our finding that bacterial load in the vagina did not predict bacterial load in either the cervix or the endometrium. Instead, it was bacterial load in the cervix that was correlated with bacterial load in the endometrium, and, while the bacterial profiles of the cervix and endometrium were similar, they were both distinct from the bacterial profile of the vagina.

Despite overall similarities in the bacterial profiles of the endometrium and the cervix, some differences were still evident. Most notably, LEfSe analyses indicated that *Lactobacillus* was more widespread and relatively abundant among cervical than endometrial samples. This finding is consistent with prior molecular studies that reported the presence of *Lactobacillus* in the cervix<sup>31,32,86,87</sup>. In the current study, the average relative abundance of *Lactobacillus* in the cervix was 19.24%. Other studies have reported a range of relative abundances of *Lactobacillus* in the cervix from 10–99%<sup>31,32,86,87</sup>. It is unclear why, in the current study, *Lactobacillus* was dominant in the vagina, common in the cervix, and yet very rare in the endometrium. Ascension of vaginal bacteria through the cervix has been suggested as a likely source of bacterial transmission to the endometrium<sup>24,88–91</sup>. For example, in an experimental study, it was shown that labeled spermatozoa-size particles (i.e. macroaggregates of human serum albumin) can translocate from the vagina to the uterus through the cervical canal within minutes in non-pregnant women<sup>92</sup>. Thus, the uterine peristaltic pump that aids in sperm transport from the cervical canal to the endometrium<sup>93</sup> might also play a role in seeding the endometrium with bacteria. However, the powerful antimicrobial activity of the cervix and cervical secretions<sup>94–111</sup> may act as an effective barrier separating the endometrium from the microbe-rich vagina. The findings of the current study suggest that the cervix may act as an effective barrier against the ascension of vaginal microbes into the endometrium. An alternative explanation is that vaginal microbes, including *Lactobacillus*, do ascend through the cervix but that their absolute and relative abundances rapidly decrease from the endocervix to the lower, middle, and upper endometrium. This may explain why *Lactobacillus* was more often detected in the endocervix/endometrium (i.e. upper genital tract) among women undergoing hysterectomies in the study performed by Mitchell *et al.*<sup>24</sup> than in the current study in which the mid-endometrium was targeted. The extent to which the vaginal and cervical microbiotas are a persistent and sustained source of bacterial populations for an endometrial microbiota, and explanations for why *Lactobacillus* immigration and colonization appears inhibited, remain to be elucidated.

The current study suggests that, if there is an endometrial microbiota, *Acinetobacter*, *Pseudomonas*, Comamonadaceae, and *Cloacibacterium* are its principal members. Chen *et al.*<sup>31</sup> previously reported that *Acinetobacter*, *Pseudomonas*, and Comamonadaceae were among the most relatively abundant bacteria within the endometrium, pouch of Douglas, and fallopian tubes, and Miles *et al.*<sup>32</sup> detected *Acinetobacter* in the cervix, endometrium, and the myometrium. Notably, *Acinetobacter* has been previously cultured from both the cervix<sup>112</sup> and the endometrium<sup>17,113,114</sup>. *Pseudomonas*<sup>17,114</sup> has been cultured from the endometrium as well.

While the exact taxonomy (i.e. species) of the prominent *Acinetobacter* OTU detected in the endometrium is unknown, the sequence analysis we conducted using BLAST showed that its consensus sequence was identical to sequences of eight *Acinetobacter* type strains that have been previously isolated from clinical samples<sup>115–121</sup>. Sequence analysis of the two prominent *Pseudomonas* OTUs in endometrial samples in this study showed that one belongs to the *P. aeruginosa* group and the other to the *P. putida* group<sup>77</sup>. The consensus sequence of the prominent *Cloacibacterium* OTU is identical to a strain of *Cloacibacterium normanense*, a species that was detected in a tissue sample of a patient with spondylodiscitis<sup>122</sup>. Lastly, the consensus sequence of the prominent Comamonadaceae OTU was identical to a strain of *Comamonas kerstersii*, a documented opportunistic pathogen that has been isolated from clinical samples<sup>123–127</sup>. Notably, these endometrial and cervical OTUs were also detected at low relative abundances in vaginal samples in this study. Similarly, in prior studies, *Acinetobacter*<sup>80,128</sup>, *Pseudomonas*<sup>80,128</sup>, and *Cloacibacterium*<sup>128</sup> have been detected in vaginal microbial communities, even when *Lactobacillus* was dominant. These taxa may therefore be common members of both the lower and upper reproductive tract. Whether the molecular signals in the current study represent a viable and residential endometrial

microbiota, and, if so, whether *Acinetobacter*, *Pseudomonas*, Comamonadaceae, and *Cloacibacterium* play a role in women's reproductive health, warrants further investigation.

It is important to note that *Acinetobacter*, *Pseudomonas*, Comamonadaceae, and *Cloacibacterium* have each been implicated as a contaminant in at least one prior molecular survey study (Supplementary Table S4). However, in the current study, the distributions and abundances of these bacterial taxa among biological samples and technical controls were not consistent with their being background DNA contaminants. For example, among the 10 blank DNA extraction kit controls we sequenced, Comamonadaceae was represented by only a single sequence read using standard PCR and four total sequence reads using touchdown PCR. Additionally, among endometrial samples, the relative abundances of Comamonadaceae were not negatively correlated with overall 16S rRNA gene abundances (i.e. bacterial load, as determined by qPCR). A negative correlation between bacterial load and OTU relative abundance is a pattern predictive of, and used by others to identify, background DNA contaminants among low biomass samples<sup>40,129</sup>. Similarly, the relative abundances of the two *Pseudomonas* OTUs and *Cloacibacterium* were also not negatively correlated with 16S rRNA gene abundances among the endometrial samples. Each of these taxa were detected in at most one-half of the background technical controls. Although *Acinetobacter* was identified among the bacterial profiles of technical controls, the relative abundances of OTU 1 (*Acinetobacter*) were positively correlated with overall 16S rRNA gene abundances (i.e., bacterial load) among endometrial samples ( $R^2 = 0.32$ ,  $p = 0.003$ ). This pattern is opposite that predicted for a background DNA contaminant. Furthermore, sequence variant analysis identified diverse *Acinetobacter*, *Pseudomonas*, Comamonadaceae, and *Cloacibacterium* sequence variants that were widespread among endometrial and cervical samples and yet were not present in background technical control samples. Together, these findings suggest that molecular signals of *Acinetobacter*, *Pseudomonas*, Comamonadaceae, and *Cloacibacterium* identified in endometrium samples are not consistent with background DNA contamination. Instead, they are suggestive of a potential resident microbiota in the endometrium whose existence must be verified through microscopy and culture.

The current study has four principal strengths. First, we limited our investigation to samples that were obtained following hysterectomy, thereby reducing the likelihood of contamination from the lower urogenital tract<sup>2,10,22</sup>. Second, we used multiple, complementary modes of inquiry: 16S rRNA gene sequencing, universal 16S rRNA gene qPCR, and *Lactobacillus*-specific targeted 16S rRNA gene qPCR. Third, we incorporated technical controls for potential background DNA contamination. Fourth, we restricted our bacterial profile analyses to samples that had both a substantive bacterial load based on the results of qPCR (i.e.  $Cq < 35$ ) and that had thorough sample coverage (i.e.  $>98\%$ ) based on 16S rRNA gene sequencing.

The current study has four primary limitations. First, all women included in the study underwent hysterectomy for gynecological disease, as endometrial biopsies from healthy individuals are rarely available. While most of the women had uterine fibroids, a condition not commonly associated with infection<sup>130</sup>, it is unclear whether the bacterial communities detected represent a normal endometrial microbiota. Second, the study was limited to the middle endometrium. Further investigation and potential delineation of a lower, middle, and upper endometrial microbiota is required. Third, swabs were not included as extraction controls. Nevertheless, there were no negative correlations between the abundances of any of the candidate endometrial microbes and sample DNA concentrations as assessed through 16S rRNA gene qPCR, a pattern illustrative of DNA contaminants. Fourth, this was a molecular microbiology study without culture and imaging components to demonstrate viability and localization of microbial communities associated with endometrial tissues<sup>129,131</sup>.

In conclusion, using complementary modes of inquiry, we detected a bacterial signature within the middle endometrium of 60% of women who underwent hysterectomy. The endometrial bacterial profiles were distinct from those of background technical controls and other body sites, except the cervix. In contrast to the vagina and cervix, *Lactobacillus* was rarely detected in the endometrium suggesting that the vaginal microbiota is not a persistent source of microbes for an endometrial microbiota. This is further supported by the finding that bacterial load in the vagina did not predict bacterial load in either the cervix or endometrium. Instead it was bacterial load in the cervix that was correlated with bacterial load in the endometrium, suggesting that, if an endometrial microbiota exists, the cervical microbiota may be a persistent source of colonizing microbes, or vice versa. The results of this study show that molecular signals from *Acinetobacter*, *Pseudomonas*, Comamonadaceae, and *Cloacibacterium*, but not *Lactobacillus*, are relatively abundant within the human mid-endometrium. The existence and viability of an endometrial microbiota must be confirmed through microscopy and culture<sup>129,131</sup>, and longitudinal studies of lower, middle, and upper endometrial microbiotas are necessary. Nevertheless, the extent to which the detected molecular signals of a potential endometrial microbiota contribute to female reproductive health and disease warrants further consideration.

## Data Availability

The MiSeq 16S rRNA gene sequence data generated during the current study have been deposited on the NCBI Sequence Read Archive (BioProject ID PRJNA543861). The raw operational taxonomic unit (OTU) data from this study are provided as Supplementary Data.

## References

- Butler, B. Value of endometrial cultures in sterility investigation. *Fertil Steril* **9**, 269–273 (1958).
- Bollinger, C. C. Bacterial flora of the nonpregnant uterus: a new culture technic. *Obstet Gynecol* **23**, 251–255 (1964).
- Mishell, D. R. Jr., Bell, J. H., Good, R. G. & Moyer, D. L. The intrauterine device: a bacteriologic study of the endometrial cavity. *Am J Obstet Gynecol* **96**, 119–126 (1966).
- Ansbacher, R., Boyson, W. A. & Morris, J. A. Sterility of the uterine cavity. *Am J Obstet Gynecol* **99**, 394–396 (1967).
- Spore, W. W., Moskal, P. A., Nakamura, R. M. & Mishell, D. R. Jr. Bacteriology of postpartum oviducts and endometrium. *Am J Obstet Gynecol* **107**, 572–577 (1970).
- Grossman, J. H. III, Adams, R. L., Hierholzer, W. J. Jr & Andriole, V. T. Endometrial and vaginal cuff bacteria recovered at elective hysterectomy during a trial of antibiotic prophylaxis. *Am J Obstet Gynecol* **130**, 312–316 (1978).

7. Pezzlo, M. T., Hesser, J. W., Morgan, T., Valter, P. J. & Thrupp, L. D. Improved laboratory efficiency and diagnostic accuracy with new double-lumen-protected swab for endometrial specimens. *J Clin Microbiol* **9**, 56–59 (1979).
8. Cooperman, N. R., Kasim, M. & Rajashekaraiyah, K. R. Clinical significance of amniotic fluid, amniotic membranes, and endometrial biopsy cultures at the time of cesarean section. *Am J Obstet Gynecol* **137**, 536–542 (1980).
9. Sparks, R. A., Purrier, B. G., Watt, P. J. & Elstein, M. Bacteriological colonisation of uterine cavity: role of tailed intrauterine contraceptive device. *Br Med J (Clin Res Ed)* **282**, 1189–1191 (1981).
10. Duff, P., Gibbs, R. S., Blanco, J. D. & St Clair, P. J. Endometrial culture techniques in puerperal patients. *Obstet Gynecol* **61**, 217–222 (1983).
11. Duff, P., Gibbs, R. S., St Clair, P. J. & Weinberg, L. C. Correlation of laboratory and clinical criteria in the prediction of postcesarean endomyometritis. *Obstet Gynecol* **63**, 781–786 (1984).
12. Awadalla, S. G., Perkins, R. P. & Mercer, L. J. Significance of endometrial cultures performed at cesarean section. *Obstet Gynecol* **68**, 220–225 (1986).
13. Eschenbach, D. A., Rosene, K., Tompkins, L. S., Watkins, H. & Gravett, M. G. Endometrial cultures obtained by a triple-lumen method from afebrile and febrile postpartum women. *J Infect Dis* **153**, 1038–1045 (1986).
14. Nelson, L. H. & Nichols, S. B. Effectiveness of the Isaacs cell sampler for endometrial cultures. *J Reprod Med* **31**, 473–477 (1986).
15. Teisala, K. Endometrial microbial flora of hysterectomy specimens. *Eur J Obstet Gynecol Reprod Biol* **26**, 151–155 (1987).
16. Møller, B. R., Kristiansen, F. V., Thorsen, P., Frost, L. & Mogensen, S. C. Sterility of the uterine cavity. *Acta Obstet Gynecol Scand* **74**, 216–219 (1995).
17. Drbohlav, P. *et al.* The effect of endometrial infection on embryo implantation in the IVF and ET program. *Ceska Gynekol* **63**, 181–185 (1998).
18. Cicinelli, E. *et al.* Prevalence of chronic endometritis in repeated unexplained implantation failure and the IVF success rate after antibiotic therapy. *Hum Reprod* **30**, 323–330 (2015).
19. Naessens, A., Foulon, W., Cammu, H., Goossens, A. & Lauwers, S. Epidemiology and pathogenesis of *Ureaplasma urealyticum* in spontaneous abortion and early preterm labor. *Acta Obstet Gynecol Scand* **66**, 513–516 (1987).
20. Cicinelli, E. *et al.* Chronic endometritis due to common bacteria is prevalent in women with recurrent miscarriage as confirmed by improved pregnancy outcome after antibiotic treatment. *Reprod Sci* **21**, 640–647 (2014).
21. Gravett, M. G., Hitti, J., Hess, D. L. & Eschenbach, D. A. Intrauterine infection and preterm delivery: evidence for activation of the fetal hypothalamic-pituitary-adrenal axis. *Am J Obstet Gynecol* **182**, 1404–1413 (2000).
22. Romero, R., Espinoza, J. & Mazor, M. Can endometrial infection/inflammation explain implantation failure, spontaneous abortion, and preterm birth after *in vitro* fertilization? *Fertil Steril* **82**, 799–804 (2004).
23. Marchesi, J. R. & Ravel, J. The vocabulary of microbiome research: a proposal. *Microbiome* **3**, 31 (2015).
24. Mitchell, C. M. *et al.* Colonization of the upper genital tract by vaginal bacterial species in nonpregnant women. *Am J Obstet Gynecol* **212**, 611.e611–619 (2015).
25. Fang, R. L. *et al.* Barcoded sequencing reveals diverse intrauterine microbiomes in patients suffering with endometrial polyps. *Am J Transl Res* **8**, 1581–1592 (2016).
26. Franasiak, J. M. *et al.* Endometrial microbiome at the time of embryo transfer: next-generation sequencing of the 16S ribosomal subunit. *J Assist Reprod Genet* **33**, 129–136 (2016).
27. Khan, K. N. *et al.* Molecular detection of intrauterine microbial colonization in women with endometriosis. *Eur J Obstet Gynecol Reprod Biol* **199**, 69–75 (2016).
28. Moreno, I. *et al.* Evidence that the endometrial microbiota has an effect on implantation success or failure. *Am J Obstet Gynecol* **215**, 684–703 (2016).
29. Verstraelen, H. *et al.* Characterisation of the human uterine microbiome in non-pregnant women through deep sequencing of the V1-2 region of the 16S rRNA gene. *PeerJ* **4**, e1602 (2016).
30. Walther-Antonio, M. R. *et al.* Potential contribution of the uterine microbiome in the development of endometrial cancer. *Genome Med* **8**, 122 (2016).
31. Chen, C. *et al.* The microbiota continuum along the female reproductive tract and its relation to uterine-related diseases. *Nat Commun* **8**, 875 (2017).
32. Miles, S. M., Hardy, B. L. & Merrell, D. S. Investigation of the microbiota of the reproductive tract in women undergoing a total hysterectomy and bilateral salpingo-oophorectomy. *Fertil Steril* **107**, 813–820.e811 (2017).
33. Tao, X. *et al.* Characterizing the endometrial microbiome by analyzing the ultra-low bacteria from embryo transfer catheter tips in IVF cycles: Next generation sequencing (NGS) analysis of the 16S ribosomal gene. *Human Microbiome Journal* **3**, 15–21 (2017).
34. Kyono, K., Hashimoto, T., Nagai, Y. & Sakuraba, Y. Analysis of endometrial microbiota by 16S ribosomal RNA gene sequencing among infertile patients: a single-center pilot study. *Reprod Med Biol* **17**, 297–306 (2018).
35. Liu, Y. *et al.* Systematic comparison of bacterial colonization of endometrial tissue and fluid samples in recurrent miscarriage patients: implications for future endometrial microbiome studies. *Clin Chem* **64**, 1743–1752 (2018).
36. Pelzer, E. S., Willner, D., Buttini, M. & Huygens, F. A role for the endometrial microbiome in dysfunctional menstrual bleeding. *Antonie Van Leeuwenhoek* **111**, 933–943 (2018).
37. Wee, B. A. *et al.* A retrospective pilot study to determine whether the reproductive tract microbiota differs between women with a history of infertility and fertile women. *Aust N Z J Obstet Gynaecol* **58**, 341–348 (2018).
38. Giudice, L. C. Challenging dogma: the endometrium has a microbiome with functional consequences! *Am J Obstet Gynecol* **215**, 682–683 (2016).
39. Baker, J. M., Chase, D. M. & Herbst-Kralovetz, M. M. Uterine microbiota: Residents, tourists, or invaders? *Front Immunol* **9**, 208 (2018).
40. Salter, S. J. *et al.* Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* **12**, 87 (2014).
41. Weiss, S. *et al.* Tracking down the sources of experimental contamination in microbiome studies. *Genome Biol* **15**, 564 (2014).
42. 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).
43. Lauder, A. P. *et al.* Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota. *Microbiome* **4**, 29 (2016).
44. Kim, D. *et al.* Optimizing methods and dodging pitfalls in microbiome research. *Microbiome* **5**, 52 (2017).
45. Dickson, R. P. *et al.* Changes in the lung microbiome following lung transplantation include the emergence of two distinct *Pseudomonas* species with distinct clinical associations. *Plos One* **9**, e97214 (2014).
46. Frank, J. A. *et al.* Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. *Appl Environ Microbiol* **74**, 2461–2470 (2008).
47. Fredricks, D. N., Fiedler, T. L., Thomas, K. K., Mitchell, C. M. & Marrazzo, J. M. Changes in vaginal bacterial concentrations with intravaginal metronidazole therapy for bacterial vaginosis as assessed by quantitative PCR. *J Clin Microbiol* **47**, 721–726 (2009).
48. Srinivasan, S. *et al.* Temporal variability of human vaginal bacteria and relationship with bacterial vaginosis. *Plos One* **5**, e10197 (2010).
49. Quinn, G. P. & Keough, M. J. Experimental design and data analysis for biologists. (Cambridge University Press, 2002).
50. Gotelli, N. J. & Ellison, A. M. A primer of Ecological Statistics. (Sinauer Associates Inc. Sunderland, MA, 2004).

51. Hammer, Ø. PAST 3 (Paleontological Statistics) version. 3.11. Home page at, <http://folk.uio.no/ohammer/past> (2015).
52. Yang, B., Wang, Y. & Qian, P. Y. Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics* **17**, 135 (2016).
53. Leung, M. H., Wilkins, D. & Lee, P. K. Insights into the pan-microbiome: skin microbial communities of Chinese individuals differ from other racial groups. *Sci Rep* **5**, 11845 (2015).
54. Fan, X., Peters, B. A., Min, D., Ahn, J. & Hayes, R. B. Comparison of the oral microbiome in mouthwash and whole saliva samples. *Plos One* **13**, e0194729 (2018).
55. Segal, L. N. *et al.* Enrichment of the lung microbiome with oral taxa is associated with lung inflammation of a Th17 phenotype. *Nat Microbiol* **1**, 16031 (2016).
56. Halfvarson, J. *et al.* Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat Microbiol* **2**, 17004 (2017).
57. Jasarevic, E., Howard, C. D., Misis, A. M., Beiting, D. P. & Bale, T. L. Stress during pregnancy alters temporal and spatial dynamics of the maternal and offspring microbiome in a sex-specific manner. *Sci Rep* **7**, 44182 (2017).
58. 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).
59. Parnell, L. A. *et al.* Microbial communities in placentas from term normal pregnancy exhibit spatially variable profiles. *Sci Rep* **7**, 11200 (2017).
60. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* **79**, 5112–5120 (2013).
61. Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K. & Mattick, J. S. ‘Touchdown’ PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res* **19**, 4008 (1991).
62. Bassis, C. M. *et al.* Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. *MBio* **6**, e00037 (2015).
63. Dickson, R. P. *et al.* Enrichment of the lung microbiome with gut bacteria in sepsis and the acute respiratory distress syndrome. *Nat Microbiol* **1**, 16113 (2016).
64. Singer, B. H. *et al.* Bacterial dissemination to the brain in sepsis. *Am J Respir Crit Care Med* **197**, 747–756 (2018).
65. Schloss, P. D. *et al.* Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**, 7537–7541 (2009).
66. Pruesse, E. *et al.* SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**, 7188–7196 (2007).
67. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**, D590–596 (2013).
68. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahe, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
69. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol* **215**, 403–410 (1990).
70. RCore Team, R. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2016 (2016).
71. Bates, D., Maechler, M., Bolker, B. & Walker, S. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–9, <http://CRAN.R-project.org/package=lme4> (2015).
72. Hothorn, T., Bretz, F. & Westfall, P. Simultaneous inference in general parametric models. *Biom J* **50**, 346–363 (2008).
73. Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral ecology* **26**, 32–46 (2001).
74. Oksanen, J. *et al.* Vegan: Community Ecology Package. R package *vegan*, version 2 (2017).
75. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol* **12**, R60 (2011).
76. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**, 581–583 (2016).
77. Anzai, Y., Kim, H., Park, J. Y., Wakabayashi, H. & Oyaizu, H. Phylogenetic affiliation of the pseudomonads based on 16S rRNA sequence. *Int J Syst Evol Microbiol* **50**, 1563–1589 (2000).
78. Moreno, I. *et al.* The diagnosis of chronic endometritis in infertile asymptomatic women: a comparative study of histology, microbial cultures, hysteroscopy, and molecular microbiology. *Am J Obstet Gynecol* **218**, 602.e601 (2018).
79. Lamont, R. F. *et al.* The vaginal microbiome: new information about genital tract flora using molecular based techniques. *Brit J Obstet Gynaecol* **118**, 533–549 (2011).
80. Ravel, J. *et al.* Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci USA* **108**(Suppl 1), 4680–4687 (2011).
81. Aagaard, K. *et al.* A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. *Plos One* **7**, e36466 (2012).
82. Petrova, M. I., Lievens, E., Malik, S., Imholz, N. & Lebeer, S. *Lactobacillus* species as biomarkers and agents that can promote various aspects of vaginal health. *Front Physiol* **6**, 81 (2015).
83. Petrova, M. I., Reid, G., Vanechoutte, M. & Lebeer, S. *Lactobacillus iners*: Friend or foe? *Trends Microbiol* **25**, 182–191 (2017).
84. Romero, R. *et al.* The vaginal microbiota of pregnant women who subsequently have spontaneous preterm labor and delivery and those with a normal delivery at term. *Microbiome* **2**, 18 (2014).
85. Greenbaum, S., Greenbaum, G., Moran-Gilad, J. & Weintraub, A. Y. Ecological dynamics of the vaginal microbiome in relation to health and disease. *Am J Obstet Gynecol* **220**, 324–335 (2018).
86. Smith, B. C. *et al.* The cervical microbiome over 7 years and a comparison of methodologies for its characterization. *Plos One* **7**, e40425 (2012).
87. Smith, B. C. *et al.* Distinct ecological niche of anal, oral, and cervical mucosal microbiomes in adolescent women. *Yale J Biol Med* **89**, 277–284 (2016).
88. Gomez, C. I., Stenback, W. A., James, A. N., Criswell, B. S. & Williams, R. P. Attachment of *Neisseria gonorrhoeae* to human sperm. Microscopical study of trypsin and iron. *Br J Vener Dis* **55**, 245–255 (1979).
89. Wolner-Hanssen, P. & Mardh, P. A. *In vitro* tests of the adherence of *Chlamydia trachomatis* to human spermatozoa. *Fertil Steril* **42**, 102–107 (1984).
90. Svenstrup, H. F., Fedder, J., Abraham-Peskir, J., Birkelund, S. & Christiansen, G. *Mycoplasma genitalium* attaches to human spermatozoa. *Hum Reprod* **18**, 2103–2109 (2003).
91. Heinonen, P. K. *et al.* Anatomic sites of upper genital tract infection. *Obstet Gynecol* **66**, 384–390 (1985).
92. Zervomanolakis, I. *et al.* Physiology of upward transport in the human female genital tract. *Ann N Y Acad Sci* **1101**, 1–20 (2007).
93. Kunz, G. *et al.* The uterine peristaltic pump. Normal and impeded sperm transport within the female genital tract. *Adv Exp Med Biol* **424**, 267–277 (1997).
94. Schumacher, G., Kim, M., Hosseinian, A. & Dupon, C. Immunoglobulins, proteinase inhibitors, albumin, and lysozyme in human cervical mucus: I. Communication: Hormonal profiles and cervical mucus changes—Methods and results. *Am J Obstet Gynecol* **129**, 629–636 (1977).
95. Hafez, E. S. E. *Human reproduction: conception and contraception*. (HarperCollins Publishers, 1980).
96. Roncalli, M., Sideri, M., Gie, P. & Servida, E. Immunophenotypic analysis of the transformation zone of human cervix. *Lab Invest* **58**, 141–149 (1988).

97. Svinarich, D. M., Wolf, N. A., Gomez, R., Gonik, B. & Romero, R. Detection of human defensin 5 in reproductive tissues. *Am J Obstet Gynecol* **176**, 470–475 (1997).
98. Quayle, A. J. *et al.* Gene expression, immunolocalization, and secretion of human defensin-5 in human female reproductive tract. *Am J Pathol* **152**, 1247–1258 (1998).
99. Valore, E. V. *et al.* Human beta-defensin-1: an antimicrobial peptide of urogenital tissues. *J Clin Invest* **101**, 1633–1642 (1998).
100. Cohen, M. Genitourinary mucosal defenses in *Sexually transmitted diseases* (ed. Holmes, K. K.) 173–190 (McGraw Hill, 1999).
101. Fichorova, R. N. & Anderson, D. J. Differential expression of immunobiological mediators by immortalized human cervical and vaginal epithelial cells. *Biol Reprod* **60**, 508–514 (1999).
102. Eggert-Kruse, W., Botz, I., Pohl, S., Rohr, G. & Strowitzki, T. Antimicrobial activity of human cervical mucus. *Hum Reprod* **15**, 778–784 (2000).
103. Fichorova, R. N., Desai, P. J., Gibson, F. C. 3rd & Genco, C. A. Distinct proinflammatory host responses to *Neisseria gonorrhoeae* infection in immortalized human cervical and vaginal epithelial cells. *Infect Immun* **69**, 5840–5848 (2001).
104. Hein, M., Valore, E. V., Helmig, R. B., Uldbjerg, N. & Ganz, T. Antimicrobial factors in the cervical mucus plug. *Am J Obstet Gynecol* **187**, 137–144 (2002).
105. Quayle, A. J. The innate and early immune response to pathogen challenge in the female genital tract and the pivotal role of epithelial cells. *J Reprod Immunol* **57**, 61–79 (2002).
106. Pudney, J., Quayle, A. J. & Anderson, D. J. Immunological microenvironments in the human vagina and cervix: mediators of cellular immunity are concentrated in the cervical transformation zone. *Biol Reprod* **73**, 1253–1263 (2005).
107. Hansen, L. K. *et al.* The cervical mucus plug inhibits, but does not block, the passage of ascending bacteria from the vagina during pregnancy. *Acta Obstet Gynecol Scand* **93**, 102–108 (2014).
108. Wira, C. R., Fahey, J. V., Rodriguez-Garcia, M., Shen, Z. & Patel, M. V. Regulation of mucosal immunity in the female reproductive tract: the role of sex hormones in immune protection against sexually transmitted pathogens. *Am J Reprod Immunol* **72**, 236–258 (2014).
109. Lee, S. K., Kim, C. J., Kim, D. J. & Kang, J. H. Immune cells in the female reproductive tract. *Immune Netw* **15**, 16–26 (2015).
110. Vornhagen, J. *et al.* Human cervical mucus plugs exhibit insufficiencies in antimicrobial activity towards Group B *Streptococcus*. *J Infect Dis* **217**, 1626–1636 (2018).
111. Elovitz, M. A. *et al.* Cervicovaginal microbiota and local immune response modulate the risk of spontaneous preterm delivery. *Nat Commun* **10**, 1305 (2019).
112. Panda, P. S., Kashyap, B. & Prasad, S. Microbiological profile of cervix of females attending *in-vitro* fertilization clinic of a tertiary care hospital, North India. *J Reprod Health Med* **2**, S7–S10 (2016).
113. Egbase, P. E., Udo, E. E., Al-Sharhan, M. & Grudzinskas, J. G. Prophylactic antibiotics and endocervical microbial inoculation of the endometrium at embryo transfer. *Lancet* **354**, 651–652 (1999).
114. Devi, C. A., Ranjani, A., Dhanasekaran, D., Thajuddin, N. & Ramanidevi, T. Surveillance of multidrug resistant bacteria pathogens from female infertility cases. *Afr J Biotechnol* **12**, 4129–4134 (2013).
115. Bouvet, P. J. & Jeanjean, S. Delineation of new proteolytic genomic species in the genus *Acinetobacter*. *Res Microbiol* **140**, 291–299 (1989).
116. Dijkshoorn, L., Van Harsselaar, B., Tjernberg, I., Bouvet, P. J. & Vaneechoutte, M. Evaluation of amplified ribosomal DNA restriction analysis for identification of *Acinetobacter* genomic species. *Syst Appl Microbiol* **21**, 33–39 (1998).
117. Karah, N. *et al.* Species identification and molecular characterization of *Acinetobacter* spp. blood culture isolates from Norway. *J Antimicrob Chemother* **66**, 738–744 (2011).
118. Nemeč, A., Dijkshoorn, L. & Jezek, P. Recognition of two novel phenons of the genus *Acinetobacter* among non-glucose-acidifying isolates from human specimens. *J Clin Microbiol* **38**, 3937–3941 (2000).
119. Nemeč, A. *et al.* Taxonomy of haemolytic and/or proteolytic strains of the genus *Acinetobacter* with the proposal of *Acinetobacter courvalinii* sp. nov. (genomic species 14 sensu Bouvet & Jeanjean), *Acinetobacter dispersus* sp. nov. (genomic species 17), *Acinetobacter modestus* sp. nov., *Acinetobacter proteolyticus* sp. nov. and *Acinetobacter vivianii* sp. nov. *Int J Syst Evol Microbiol* **66**, 1673–1685 (2016).
120. Turton, J. F., Shah, J., Ozongwu, C. & Pike, R. Incidence of *Acinetobacter* species other than *A. baumannii* among clinical isolates of *Acinetobacter*: evidence for emerging species. *J Clin Microbiol* **48**, 1445–1449 (2010).
121. Vaneechoutte, M. *et al.* Identification of *Acinetobacter* genomic species by amplified ribosomal DNA restriction analysis. *J Clin Microbiol* **33**, 11–15 (1995).
122. Stavnsbjerg, C., Frimodt-Møller, N., Moser, C. & Bjarnsholt, T. Comparison of two commercial broad-range PCR and sequencing assays for identification of bacteria in culture-negative clinical samples. *BMC Infect Dis* **17**, 233 (2017).
123. Almuzara, M. *et al.* Unusual presentations of *Comamonas kerstersii* infection. *New Microbes New Infect* **19**, 91–95 (2017).
124. Almuzara, M. N. *et al.* Intra-abdominal infections due to *Comamonas kerstersii*. *J Clin Microbiol* **51**, 1998–2000 (2013).
125. Biswas, J. S., Fitchett, J. & O'Hara, G. *Comamonas kerstersii* and the perforated appendix. *J Clin Microbiol* **52**, 3134 (2014).
126. Opota, O. *et al.* Bacteremia caused by *Comamonas kerstersii* in a patient with diverticulosis. *J Clin Microbiol* **52**, 1009–1012 (2014).
127. Zhou, Y. H., Ma, H. X., Dong, Z. Y. & Shen, M. H. *Comamonas kerstersii* bacteremia in a patient with acute perforated appendicitis: A rare case report. *Medicine (Baltimore)* **97**, e9296 (2018).
128. Yeoman, C. J. *et al.* A multi-omic systems-based approach reveals metabolic markers of bacterial vaginosis and insight into the disease. *Plos One* **8**, e56111 (2013).
129. de Goffau, M. C. *et al.* Recognizing the reagent microbiome. *Nat Microbiol* **3**, 851–853 (2018).
130. Laughlin, S. K., Schroeder, J. C. & Baird, D. D. New directions in the epidemiology of uterine fibroids. *Semin Reprod Med* **28**, 204–217 (2010).
131. Theis, K. R. *et al.* Does the human placenta delivered at term have a microbiota? Results of cultivation, quantitative real-time PCR, 16S rRNA gene sequencing, and metagenomics. *Am J Obstet Gynecol* **220**, 267.e261 (2019).

## Acknowledgements

We gratefully acknowledge the patients whose gynecological experiences were the basis of this study. We thank members of the Perinatology Research Branch and the University Hospital of Padua staff for their respective contributions to the development and publication of this article. This research was supported, in part, by the Perinatology Research Branch, Division of Obstetrics and Maternal-Fetal Medicine, Division of Intramural Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, U.S. Department of Health and Human Services (NICHD/NIH/DHHS); and, in part, with Federal funds from NICHD/NIH/DHHS under Contract No. HHSN275201300006C. It was further supported by the Wayne State University Perinatal Research Initiative in Maternal, Perinatal and Child Health. Dr. Romero has contributed to this work as part of his official duties as an employee of the United States Federal Government.

### Author Contributions

A.D.W., R.R., M.T.G., N.G.L. and K.R.T. conceived and designed the study. A.D.W., R.R., M.T.G., N.G.L., M.R.T., V.G.F. and K.R.T. acquired and analyzed data. A.D.W., R.R., N.G.L., P.P., E.J., S.H., C.D.H. and K.R.T. interpreted the data. All authors drafted the manuscript.

### Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-46173-0>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019