Microbiome of the Lower Genital Tract in Chinese Women with Endometriosis by 16s rRNA Sequencing Technique: A Pilot Study

Sikai Chen, Zhiyue Gu, Wen Zhang, Shuangzheng Jia, Yushi Wu, Ping Zheng, Yi Dai, Jinhua Leng*

Department affiliation and address: Department of Obstetrics and Gynecology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College; No. 1 Shuaifuyuan, Dongcheng District, Beijing 100730, China

***Corresponding Author:** Jinhua Leng

Email addresses of corresponding authors:

*JinHua Leng: lengjenny@vip.sina.com

Abstract

Objective: Endometriosis is a chronic disease characterized by the growth of endometrial cells outside the uterine cavity. The dysfunction of the immune system is strongly associated with the progression of endometriosis, and is also correlated to the diversity of microbiota in the genital tract. According to previous studies, the microbiota significantly contributes to multi-systemic function, but the evidence of endometriosis and adenomyosis remains insufficient. Thus, the present study attempted to identify the characteristics of microbiota in endometriosis patients, and the connection between microbiota and immunological dysfunction.

Methods: In order to compare and explore the potential microbiota correlated to endometriosis and adenomyosis in the genital tract, 134 samples obtained from the cervical canal, posterior fornix and uterine cavity were analyzed by 16s-rRNA sequencing. The raw data was filtered, analyzed and visualized, and bio-information

methods were used to identify the different and distinctive characteristics of microbiota. Results: Two different locations near the cervix, cervical canal and posterior fornix, exhibited no differences in alpha diversity. Among the different disease groups, five microbiota were distinctive in the genus level, and *Atopobium* presented with the greatest significance in adenomyoisis-endometriosis patients. The LEfSe analysis failed to identify the special biomarkers, while several characteristic functions were identified through PICRUSt.

Conclusion: *Lactobacillus* is dominant in the female lower genital tract, and *Atopobium* is distinctively higher in patients with endometriosis combined with adenomyosis. Several different functions of microbiota were explored, and these are found to be associated with endometriosis and adenomyosis. These findings may provide a new concept of microbiota/immune system/endometriosis system. There is an urgent need to investigate the potential microbial biomarkers of endometriosis in the future.

Trial registration: The Ministry of Science and Technology of the People's Republic of China, National Program on Key Basic Research Project of China (SQ2017YFSF080001), and The Institutional Review Board at Peking Union Medical College Hospital approved the present study (approval No. JS-1532).

Keywords: endometriosis; microbiome; 16s-rRNA sequencing; *Atopobium*

Introduction

Endometriosis is known as a benign estrogen-depend gynecological disease that mainly affects women in reproductive age, and prevails in 2-10% of the population [1]. It is characterized by the growth of endometrial glands and stromal cells inside and outside the pelvic cavity [2-4]. The symptom of endometriosis mainly presents with infertility, dysmenorrhea, chronic pelvic pain and dyspareunia [5], and it also places a

 doi:10.20944/preprints201911.0147.v1

large burden on the health care system [4]. There are three main types of endometriosis: ovarian endometriosis, deep infiltrating endometriosis (DIE), and peritoneal endometriosis. The other special types of endometriosis merely account for a small portion, but still significantly contributes to low quality of life. DIE contributes the most in pelvic pain syndrome, while ovarian endometriosis mainly presents with a pelvic mass [6]. Although retrograde menstruation is the most acceptable theory at present, the etiology of endometriosis remains unclear after many researches for over a century. Oxidative stress and genetic features are the other possible important factors that lead to endometriosis [7], and these may affect the immune system and contribute to the growth of endometrial tissue outside the uterine cavity. Many different immunological cells are associated with endometriosis [2], which include macrophages [8], mast cells [9], neutrophils, dendritic cells and gliocytes [2]. Some immunological factors, such as interleukins (IL), interferons (IFN), tumor necrosis factors (TNF) and macrophage inflammatory protein (MCP) [7, 10], are also correlated to endometriosis [2]. Nevertheless, endometriosis is not only a regional restricted disease, but also a systemic immunological dysfunction disease. In clinic, it remains as a challenge to clearly diagnose endometriosis through only non-invasive methods. Ultrasound, magnetic resonance image (MRI) and CA125 blood test are the recommended methods to evaluate the disease [7], while laparoscopic surgery and pathological examination are the gold standards of diagnosis. To date, gynecologists are short of methods to evaluate the munity of endometriosis in the normal population. Hence, it is imperative to establish a non-invasive diagnostic model to evaluate and carry out early interventions.

According to a previous study, immunological and genetic function disturbance may increase the munity of bacterial vaginosis (BV) through interleukin and CD4 cell

 doi:10.20944/preprints201911.0147.v1

overexpression [11]. BV is also associated with higher numbers of CCR5 cells, the CD4/CD8 rate, and the IL-1β and TNF-α level in vaginal-cervical mucosa [12]. In healthy women, the vaginal microbiota is dominated by *Lactobacillus*, and the low-pH environment is maintained by secreting lactate [13, 14]. The dysfunction of immunological factors not only increase the risks of HPV/HIV infection, but also increase the risk of bacterial infection [13, 15]. Sufficient evidences are available on the micro-environment in the cervical-vaginal region, but it remains debatable whether the uterine cavity is sterilized. Khan *et.al* validated that lipopolysaccharide is detectable in the endometrium, and that it is highly correlated to endometriosis [16]. Other researchers have also successfully verified that the uterine cavity is non-sterilized [17, 18] through 16S ribosomal-RNA gene sequencing and bacterial culture [19]. However, it remains uncertain whether these positive results were caused by accident contamination from the cervical-vaginal canal during sample collection, or by some unknown etiologies, such as pregnancy [20] and sub-clinical infection [18, 19].

A 'second human genome project' was proposed in 2001 to explore the human microbiome composition, and investigate the interactions between the microbiome and host [21]. Two methods were mainly used to analyze the microbiome composition, which include culture-based technology and sequencing-based technology. However, the low throughput culture-based method could not detect the whole microbes of one specific site. The 16S ribosomal-RNA gene sequencing, which is a kind of next-generation high throughput DNA sequencing technology, could make up the deficiency of the culture-based method, and has become the specific standard approach to identity the microbiome [22].

To our knowledge, few relevant studies that simply mentioned the different microbiome composition of the reproductive duct between patients with or without

 doi:10.20944/preprints201911.0147.v1

endometriosis by performing 16S rRNA gene amplicon sequencing [17, 18] were found in the systematic review. The present study aimed to establish a predictive method for the early and non-invasive diagnosis of endometriosis through the 16s-rRNA gene amplicon sequencing technique.

Materials and Methods

Patients and sample collection

The present study was approved by the institutional review board at Peking Union Medical College Hospital. The outpatients and inpatients, who were ready to undergo pre-operational preparation in Peking Union Medical College Hospital from April 2018 to February 2019, were included in the present study. The inclusion criteria included the following items: patients within 18-45 years old, and patients with regular menstrual cycle (28 ± 7 days), non-antibiotics exposure within 30 days, no sexual activity within 48 hours, no douching and vaginal medications within five days, and no cervical treatment within one week. The exclusion criteria included the following items: patients with BV, cervical inflammation, pelvic inflammatory disease, any acute inflammation, cancer and autoimmune disorders, pregnant patients, patients with intra-uterine devices, and patients who were in menstrual period at the time of sampling. An informed consent was obtained from each participant. Samples were collected from the cervix on the day of clinical visit for outpatients or before pre-operational vaginal douching for inpatients. Disposable swabs (Jiangsu Tianli Medical Instrument Co., Ltd.) were used for sampling the cervical canal (A) and posterior fornix (B), and vacuum suck tubes (Jiangsu Tianli Medical Instrument Co., Ltd., Guardking, JDC-II) were used for sampling the endometrium (C). In order to minimize the contamination from the cervical canal and vagina, the cervical canal was sterilized with iodine, and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2019 doi:10.20944/preprints201911.0147.v1

the vacuum suck tube was inserted into the uterine cavity to avoid any contact with the vaginal wall. Then, the samples were placed in ice, stored at -80°C, and transported on dry ice to Annoroad Gene Technology Co. Ltd. (Beijing, China) for further analysis.

The 16S ribosomal-RNA gene sequencing

Total genome DNA was extracted from the samples using the CTAB/SDS method. The DNA concentration and purity were monitored on 1% agarose gels. According to the concentration, the DNA was diluted to 1 ng/μL using sterile water. Then, the 16S rRNA genes of distinct regions (16SV3-V4) were amplified using specific primers 338F (5'-ACTCCTACGGGAGGCAGCAG-3') and 806R (5'-GGACTACHVGGGTWTCTAAT-3') with the barcode. All PCR reactions were carried out with the Phusion® High-Fidelity PCR Master Mix (New England Biolabs). A certain volume of the 1X loading buffer (containing SYBR green) was mixed with the PCR products, and electrophoresis was performed on 2% agarose gel for detection. Samples with a bright main strip between 400-450 bp were chosen for further experiments. The PCR products were mixed in equidensity ratios. Then, the mixture of PCR products was purified using the Qiagen Gel Extraction Kit (Qiagen, Germany). The sequencing libraries were generated using a TruSeq® DNA PCR-Free Sample Preparation Kit (Illumina, USA), according to manufacturer's recommendations, and index codes were added. The library quality was assessed on the Qubit@ 2.0 Fluorometer (Thermo Scientific) and Agilent Bioanalyzer 2100 system. Lastly, the library was sequenced on the lluminaHiSeq2500 platform, and 250 bp paired-end reads were generated.

The 16s rRNA sequence analysis

 doi:10.20944/preprints201911.0147.v1

The raw sequence reads of the 16s rRNA gene sequences were analyzed and quality filtered using the vsearch and usearch tool [23, 24]. The amplicon sequence variants (ASV) were taxonomically classified through Unoise3 [25], and using the Greengenes 16S rRNA gene reference database. The taxonomic composition of microbial communities was visualized using R and STAMP v.2.1.3. The community clustering was measured by unweighted uniFrac and weighted unifrac distance, based on the normalized Operational Taxonomy Units (OTU) table.

LDA Effect Size (LEfSe) analysis

In order to identify the distinct cervical canal microbiota between different groups, the LEfSe method was used to compare the composition of the cervical canal microbiota using an online tool (www.ehbio.com). After the pairwise comparison, no discriminative species were found among the four groups.

Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt)

After filtering and the analysis of OTUs, PICRUSt [25] was performed to determine the different functions of the vaginal bacteria. The vsearch [23] tool was used to analyze the four different groups of participants, and the KEGG Orthology (KO3) functional profiling of the vaginal microbiota was performed. The matrix was normalized by dividing the absolute amount of each functional gene by the total number of reads assigned to the functional genes in each sample (each sample was normalized for 3,900). Merely statistically significant results ($P<0.05$) were presented in the figures using STAMP v2.1.3.

 doi:10.20944/preprints201911.0147.v1

Results

Samples and participant characteristics

A total of 68 patients were included in the present study, cervical canal (A) and posterior fornix (B) samples were collected from each participant. Among these samples, 67 samples obtained from the cervical canal (A) and 65 samples obtained from the posterior fornix (B) successfully proceeded to the PCR. Due to the informed consent obtained from each patient, merely two samples were collected from the uterine cavity (C) and successfully proceeded to the PCR. Hence, a total of 134 samples were available. Since 15 out-patients did not receive surgical interventions, their final diagnosis were obtained from the radiological examinations, which included their trans-vaginal ultrasound and magnetic resonance imaging (MRI). A total of 55 participants underwent surgical intervention, and their medical records, including surgery records and pathological reports, were available. Furthermore, 19 (27.94%) participants were diagnosed with adenomyosis, 20 (29.41%) participants were diagnosed with endometriosis (including ovarian endometriosis, DIE, peritoneal type and other special types), seven (10.29%) participants were diagnosed with adenomyosis accompanied with endometriosis, and 36 (52.94%) patients were diagnosed with classified as control group, which included infertility, leiomyoma, ovarian borderline tumor and teratoma. Among the 134 samples, 25 samples were assigned to the endometriosis (EM) group, 67 samples were assigned to the control (CT) group, 14 samples were assigned to the adenomyosis-endometriosis (AMEM) group, and 28 samples were assigned to the adenomyosis (AM) group.

The 16S rRNA gene sequencing of the eligible 67 cervical canal samples, 65 posterior fornix samples and two endometrial samples yielded a total of 7.35 million raw sequences, with 30,084-173,739 sequences per sample, and the average length of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2019 doi:10.20944/preprints201911.0147.v1

sequence for each sample ranged within 417.21-499.99 bp (Supplementary data 1). According to the alpha rarefaction curve estimation, the number of sequences can well represent the microbial diversity of each community (Supplementary data 2).

Microbiota differences in the two different sites near the cervix

In order to evaluate the different microbiota in different locations in the cervix, Principal Component Analysis (PCA) was conducted. Since endometriosis is a disease that closely depends on an endometrial environment, determining whether the endometrium is sterilized remains debatable, and there is a risk for the endometrial samples to be contaminated. The investigators attempted to determine whether the other sites near the cervix could take the place of the microbiota in the uterine cavity. Two samples from the endometrium were not enough to be analyzed in PCA plot. Hence, two locations near the cervix were analyzed, as previously mentioned (cervical canal and posterior fornix). All three levels (order, family and genus) of microbiota revealed no significance in the PCA plot (Fig. 1).

Microbiota composition of different diseases

In order to determine whether the composition of microbiota differ, an alpha diversity analysis was conducted. The alpha diversity was computed in four different methods, including the Chao index, Richness index, Shanoon_e index and Simpson index. Four groups of boxplots were generated, and no statistical significance was found in alpha diversity (Fig. 2). This result suggests that the number and proportion of bacterial type may not be correlated to these diseases.

In addition, a beta diversity analysis was also performed to determine the proportion of each OTU, and determine whether the differences were significant. As

 doi:10.20944/preprints201911.0147.v1

shown in Figure 3, the analysis was conducted in three different levels. *Lactobacillus* was the dominant genus in the vagina in all three levels, as previously reported. For patients with adenomyosis-endometriosis, in terms of order level, *Coriobacteriales* shared the largest proportion in the four groups. *Coriobacteriaceae* was more dominant than any of the other three groups in the family level. In the genus level, *Atopobium* was greater than any of the other groups.

After filtering the raw data, the microbiota with an OTU abundance of >1% was selected in all four groups of participants. Six kinds of genus of microbiotas were filtered, and these were presented in the circular plot (Fig. 4). *Lactobacillus* was the most prevalent microbiota, and the four groups of participants shared this, which was almost the same. However, a large proportion of *Atopobium* was taken by the AMEM group, while a small proportion was occupied by the other three groups.

Principal co-ordinates analysis (PCoA) is a method that could simplify the distances of samples from a multiple-dimensional to two-dimensional figure [27]. The investigators attempted to determine the differences among the four groups of participants from the PCoA. The weighted-unifrac and unweight-unifrac were analyzed, as shown in **Figures 5** and **6**. However, the distribution of these four disease groups were not fully separated.

Visualized bar plots were used to explore the different abundance of microbiota with statistical significance ($P<0.05$). The OTU data were analyzed in the genus and family levels, hoping to determine the quantitative differences among the four groups of participants (Fig. 7). In the genus level, four genus of microbiota revealed a higher abundance in the AMEM group, when compared to any of the other three groups. Furthermore, *Atopobium, Campylobacter, Ezakiella, Faecalibaterium* and *Escherichia/Shigella* exhibited a higher abundance in the AMEM group, when

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2019 doi:10.20944/preprints201911.0147.v1

compared to the CT group and EM group. In the family level, merely *Coriobacteriaceae* and *Campylobacteriaceae* was significantly higher in AMEM group, when compared to any of the other three groups.

Phylogenetic Investigation of Communities by Reconstruction of Unobserved States

After the analysis of the different compositions of microbiota among the four groups of diseases, another important question is why does these significant differences exists, and does these contribute to some unknown etiologies of diseases? PICRUSt was performed to explore the different functions of different microbiota. Furthermore, KO3 functional profiling was performed in two-two comparison (Supplementary Fig. 1).

In the AMEM and EM groups (Supplementary Fig. 1a), 65 KO3 functions were labeled with statistically significance (P<0.05), and nearly half of these were in a small portion of less than 1%. Hence, it is hard to explain the real impact and significance to the diseases. In addition, several items did show great importance, according to the results: *environmental information processing/signal transduction/two-component system, genetic information processing/transcription/transcription factor*, and *environmental information processing/membrane transport/secreting system*. This was lesser in the AMEM group than in the EM group in such items. In contrast, this was greater in the AMEM group than in the EM groups in terms of *genetic information processing/translation/translation factor, metabolism/amino acid metabolism/amino acid related enzyme, genetic information processing/folding sorting, degradation/protein export, genetic information processing/translation/ribosome, unclassified/genetic information processing/translation protein, metabolism/nucleotide*

 doi:10.20944/preprints201911.0147.v1

metabolism/purine metabolism, and *metabolism/glycan biosynthesis and metabolism/peptidoglycan biosynthesis*.

In the AMEM and CT groups (Supplementary Fig. 1b), 60 KO3 functions were identified (P<0.05). *Genetic information processing/translation/ribosome, metabolism/amino acid metabolism/amino acid related enzyme, unclassified/genetic information processing/translation protein, genetic information processing/translation/Aminoacyl-tRNA biosynthesis*, and *metabolism/energy metabolism/carbon fixation in photosynthetic organism* exhibited a much higher proportion in the AMEM group, when compared to the CT group. Meanwhile, *environmental information processing/signal transduction/two-component system, metabolism/carbohydrate metabolism/pyruvate metabolism*, and *genetic information processing/transcription/transcription factor* shared a greater proportion in the CT group, when compared to the AMEM group.

In the AMEM and AM groups (Supplementary Fig. 1c), 24 functions with statistical significance were identified. This was significantly higher the AM group than in the AMEM group in terms of *environmental information processing/signal transduction/two-component system*, *metabolism/amino acid metabolism/cysteine*, and *methionine metabolism* and *metabolism/amino acid metabolism/lysine biosynthesis*. Furthermore, this was higher in the AMEM group than in the AM group in terms of *metabolism/glycan biosynthesis and metabolism/peptidoglycan biosynthesis*, *genetic information processing/translation/ribosome*, *metabolism/energy metabolism/carbon fixation in photosynthetic organism*, and *metabolism/amino acid metabolism/amino acid related enzyme*.

In the AM and CT groups (Supplementary Fig. 1d), 10 KO3 functions were labeled. However, merely one function exhibited a great significance: *metabolism/energy*

 doi:10.20944/preprints201911.0147.v1

metabolism/methane metabolism. This was higher in the AM group in this KO3 function.

In the CT and EM groups (Supplementary Fig. 1e), 20 KO3 functions were identified. *Metabolism/metabolism of cofactors and vitamins/folate biosynthesis*, and *Metabolism/metabolism of cofactors and vitamins/pantothenate and CoA synthesis* shared a higher proportion in the CT group, while *metabolism/lipid metabolism/synthesis and degradation of ketone bodies* was higher in the EM group.

Finally, in the AM and EM groups (Supplementary Fig. 1f), merely *metabolism/lipid metabolism/synthesis and degradation of ketone bodies* was identified, and this was significantly higher in the EM group than in the AM group.

Overall, several functions overlapped in the two-two comparison. The AMEM group had a greater proportion of *environmental information processing/signal transduction/two-component system, genetic information processing/transcription/transcription factor*, *genetic information processing/translation/ribosome*, and *metabolism/glycan biosynthesis and metabolism/peptidoglycan biosynthesis*. However, the EM group had a higher proportion of *Metabolism/lipid metabolism/synthesis and degradation of ketone bodies.*

Discussion

Endometriosis is a disease that involves endometrium tissues and immune system dysfunction [2]. First, the investigators attempted to determine the characteristics of the microbiota in the uterine cavity. After the collection of samples from the uterine cavities, the outcome of the PCR was barely satisfactory, when compared to samples obtained from the cervix. Furthermore, no valid evidence demonstrated that the uterine cavity is sterilized, since the immune system can still affect the micro-environment in

 doi:10.20944/preprints201911.0147.v1

the uterine cavity [8, 16]. Some researchers have successfully proven the non-sterilized environment through 16s-rRNA sequencing [17] and bacterial culture [16] methods. However, it remains uncertain whether other unknown sub-clinical infections or other underlying diseases can affect the environment in the uterine cavity. The "bacterial contamination hypothesis" was proposed by K Khan *et al.* [16]. This hypothesis suggests that the lipopolysaccharide level of *Escherichia coli* is significantly higher in endometriosis patients, and that this significantly contributes to the progression of endometriosis through the dysfunction of the immune system. In addition, A Takebayashi *et al.* suggested that endometriosis is strongly associated with endometritis [28, 29]. From another aspect, the mucus of the vagina and cervix are enriched with bacteria, and it is possible for lipopolysaccharides or bacterial reflux to be present in the uterine cavity and cause immunological reactions. In the present study, two different sites near the cervix (cervical canal A and posterior fornix B) were compared, and an attempted was made to identify a marker site that was able to replace the endometrium environment. However, the available samples obtained from the uterine cavity was not enough, and it was difficult to compare the OTUs with the cervix samples. Furthermore, after the comparison of microbiota between A and B, no differences were obtained through the PCA plots. This result suggests that the microbiota obtained from different locations near the cervix was not a distinctive marker site, and that immunological dysfunction may affect the vaginal and uterine cavity micro-environment. However, the determination of the closeness of their association remains unclear.

The association of adenomyosis and endometriosis had not been clearly demonstrated [30]. At present, there are no studies on 16s-rRNA sequencing in adenomyosis. Hence, adenomyosis patients were included, and an attempted was made to determine the underlying associations through 16s-rRNA sequencing. Four types of

 doi:10.20944/preprints201911.0147.v1

index in alpha diversity was conducted: Chao index, Richness index, Shanoon_e index and Simpson index. However, no statistical significance was found. This result suggests that endometriosis and adenomyosis may not differ in the composition of microbiota. In addition, all four groups of participants exhibited no significances in alpha diversity. Furthermore, the numbers of existing OTUs were stable, and the composition of the microbiota has an unclear impact to the disease, according to the present results. Hence, more cases should be included and more studies should be conducted to establish a better theory of microbiota relationship. Fortunately, the beta diversity revealed a possible result that adenomyosis-endometriosis patients may be correlated to *Atopobium*, and that *Atopobium* was more abundant in adenomyosis-endometriosis patients than in any other groups. *Atopobium* has been previously reported as a BV-related microbiota [31], and that it also colonizes in healthy women [32]. Some sub-clinical infections may still exist, although BV was excluded in these present participants through verbal confirmation and microbial nuclei acid detection (*Candida Albican*, *Trichomonas vaginalis* and *Gardnerella*). Previous studies have also demonstrated that endometritis is strongly correlated to endometriosis [28, 29]. The "bacterial contamination hypothesis" demonstrated that lipopolysaccharide is detected in the endometrium, and activates the LPS/TLR4 cascade in women with endometriosis [16]. *Streptococcaceae, Staphylococaceae* and *Enterobacteriaceae* were significantly elevated in gonadotropin-releasing hormone agonist (GnRHa)-treated women with endometriosis [17, 19]. These evidences may provide some information that endometriosis is possibly correlated to bacterial contamination, and the present study supports this hypothesis, to a certain extent. However, it remains to be determined which kind of bacteria contributes the most to endometriosis, and how these infect the genital tract. According to the present results, *Atopobium* was statistically significant in

adenomyosis-endometriosis patients, which was opposite to the result reported by B Ata *et al.* [33] *Atopobium* is carcinogenic, and it can facilitate infection through *Porphyromonas* species, which can intracellularly manifest and disrupt cell regulatory functions, leading to a carcinogenic trigger [33]. Furthermore, it also relates to higher IL-1β levels [13]. A previous study also revealed that high IL-1β levels are associated with higher risk of endometriosis [34]. It was assumed that adenomyosis and endometriosis may correlate to *Atopobium* through IL-1β. However, it remains to be determined how the relationship of adenomyosis and endometriosis in the microbiome pathway can be explain through 16s-rRNA sequencing. the present study suggests that *Atopobium* may be involved in the progression of endometriosis through certain unknown mechanisms, such as IL-1β and *Atopobium* infection.

PICRUSt revealed a novel concept that endometriosis and adenomyosis may be associated with certain mechanisms through a predictive microbiota function analysis. *Ribosome* is a function that should be noticed, and previous study on ribosomes suggested that the upregulation of ribosomes may promote the progression of endometriosis [35]. However, there are no reports on the association of adenomyosis and ribosomes, to date. The *two-component system* is another distinctive function, and it involves the regulation of gene expression in response to environmental signals, such as antibiotic exposure in Gram-negative bacteria [36]. However, further explorations are needed to determine whether antibiotic resistance is a trigger of endometriosis. It has been reported that the *transcription factor* of the endometriosis is involved the regulatory program of endometriosis [37], and can act as a therapeutic target [38]. Furthermore, the *peptidoglycan biosynthesis* function is significantly higher in adenomyosis-endometriosis, and the cell wall of bacteria consists of peptidoglycan and contributes to antibiotic resistance [39]. However, there are no reports on the

 doi:10.20944/preprints201911.0147.v1

association between peptidoglycan and endometriosis at present.

The present study demonstrated that five kinds of genus of microbiota were distinctive, with an OTU abundance of >1%. This should be noticed in future studies. The present study also has some limitations: (1) The PCR quality of the intra-uterine samples was not satisfactory. Thus, the data was not enough to conduct the analysis, and merely samples obtained from the cervical canal and posterior fornix could be compared. (2) The LEfSe analysis did not reveal any reasonable biomarkers. However, five microbiotas revealed a statistical significance in the genus level. However, it remains to be determined whether *Atopobium* and the four other microbiotas can be used as a biomarker. (3) Sub-clinical infection and sub-clinical immunological dysfunction is hard to detect in these participants. Hence, it remains unknown how these disorders may affect the result, and whether these disorders are the underlying causes of endometriosis and adenomyosis. (4) Although a PICRUSt predictive analysis was conducted in the KEGG database, the real functional connection between microbiota and the human body remains to be determined. Further functional studies and researches are needed.

Conclusion

The present study suggests that the microbiota of the cervix is distinctive in adenomyosis-endometriosis patients, and that *Lactobacillus* remains dominant. The microbiota in endometriosis and adenomyosis women do not differ in alpha diversity, but the five types of microbiota were distinctive in the genus level, while *Atopobium* shared the largest proportion in adenomyosis-endometriosis patients. *Two-component system, transcription factor, ribosome, peptidoglycan biosynthesis* and *synthesis and degradation of ketone bodies* were possibly the distinctive functions in endometriosis

and adenomyosis. These findings may provide a novel view of the microbiota-immune function-endometriosis system, in which the immune system is associated with the microbiota and endometriosis. However, further studies are needed to understand such network. Furthermore, the present pilot study provides a novel concept that bacteria in lower genital tract may be correlated to the etiology of endometriosis, the bacterial features of endometriosis may become a non-invasive diagnostic method, and even bacterial therapy could be a possible method to treat endometriosis in the future.

References:

1. Dai Y, Zhou Y, Zhang X, Xue M, Sun P, Leng J, Chapron C: **Factors associated with deep infiltrating endometriosis versus ovarian endometrioma in China: a subgroup analysis from the FEELING study.** *BMC Women's Health* 2018, **18:**1-9.

2. Ahn SH, Monsanto SP, Miller C, Singh SS, Thomas R, Tayade C: **Pathophysiology and Immune Dysfunction in Endometriosis.** *Biomed Res Int* 2015, **2015:**1-12.

3. Yuan M, Li D, Zhang Z, Sun H, An M, Wang G: **Endometriosis induces gut microbiota alterations in mice.** *Hum Reprod* 2018, **4:**607-616.

4. McKinnon BD, Bertschi D, Bersinger NA, Mueller MD: **Inflammation and nerve fiber interaction in endometriotic pain.** *Trends Endocrinol Metab* 2015, **26:**1-10.

5. Dai Y, Zhang X, Xue M, Zhou Y, Sun P, Leng J: **Not Having Been Breastfed May Protect Chinese Women From Developing Deep Infiltrating Endometriosis: Results From Subgroup Analyses of the FEELING Study.** *Reprod Sci* 2019**:**288109685.

449 6. Khan KN, Kitajima M, Fujishita A, Hiraki K, Matsumoto A, Nakashima M,
450 Masuzaki H: **Pelvic pain in women with ovarian endometrioma is mostly associated**
451 **with coexisting peritoneal lesions.** *Hum Reprod* 2012, **28:**109-118.
452 7. Dai Y, Li X, Shi J, Leng J: **A review of the risk factors, genetics and treatment**
453 **of endometriosis in Chinese women: a comparative update.** *Reprod Health* 2018,
454 **15**.
455 8. Wu J, Xie H, Yao S, Liang Y: **Macrophage and nerve interaction in**
456 **endometriosis.** *J Neuroinflamm* 2017, **14:**1-11.
457 9. D'Cruz OJ, Uckun FM: **Targeting mast cells in endometriosis with janus kinase**
458 **3 inhibitor, JANEX-1.** *Am J Reprod Immunol* 2007, **58:**75-97.
459 10. Ellis A, Bennett DLH: **Neuroinflammation and the generation of neuropathic**
460 **pain.** *Brit J Anaesth* 2013, **111:**26-37.
461 11. Murphy K, Mitchell CM: **The Interplay of Host Immunity, Environment and**
462 **the Risk of Bacterial Vaginosis and Associated Reproductive Health Outcomes.** *J*
463 *Infect Dis* 2016, **214 Suppl 1:**S29-S35.
464 12. Thurman AR, Kimble T, Herold B, Mesquita PM, Fichorova RN, Dawood HY,
465 Fashemi T, Chandra N, Rabe L, Cunningham TD, et al: **Bacterial Vaginosis and**
466 **Subclinical Markers of Genital Tract Inflammation and Mucosal Immunity.** *AIDS*
467 *Res Hum Retroviruses* 2015, **31:**1139-1152.
468 13. Mitchell C, Marrazzo J: **Bacterial Vaginosis and the Cervicovaginal Immune**
469 **Response.** *Am J Reprod Immunol* 2014, **71:**555-563.
470 14. Boskey ER, Cone RA, Whaley KJ, Moench TR: **Origins of vaginal acidity: high**

 doi:10.20944/preprints201911.0147.v1

D/L lactate ratio is consistent with bacteria being the primary source. *Hum Reprod* 2001, **16:**1809-1813.

15. Wijgert JHHM, Borgdorff H, Verhelst R, Crucitti T, Francis S, Verstraelen H, Jespers V: **The Vaginal Microbiota: What Have We Learned after a Decade of Molecular Characterization.** *Plos One* 2014, **8:**e105998.

16. Khan KN, Fujishita A, Hiraki K, Kitajima M, Nakashima M, Fushiki S, Kitawaki J: **Bacterial contamination hypothesis: a new concept in endometriosis.** *Reprod Med Biol* 2018, **17:**125-133.

17. Khan KN, Fujishita A, Masumoto H, Muto H, Kitajima M, Masuzaki H, Kitawaki J: **Molecular detection of intrauterine microbial colonization in women with endometriosis.** *Eur J Obstet Gyn R B* 2016, **199:**69-75.

18. Chen C, Song X, Wei W, Zhong H, Dai J, Lan Z, Li F, Yu X, Feng Q, Wang Z, et al: **The microbiota continuum along the female reproductive tract and its relation to uterine-related diseases.** *Nat Commun* 2017, **8**.

19. Khan KN, Fujishita A, Kitajima M, Hiraki K, Nakashima M, Masuzaki H: **Intra-uterine microbial colonization and occurrence of endometritis in women with endometriosis†.** *Hum Reprod* 2014, **29:**2446-2456.

20. Aagaard K, Ma J, Antony KM, Ganu R, Petrosino J, Versalovic J: **The Placenta Harbors a Unique Microbiome.** *Sci Transl Med* 2014, **6:**237r-265r.

21. Wolfe AJ, Brubaker L: **"Sterile Urine" and the Presence of Bacteria.** *Eur Urol* 2015, **68:**173-174.

22. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh

493 PJ, Fierer N, Knight R: **Global patterns of 16S rRNA diversity at a depth of millions**
494 **of sequences per sample.** *Proc Natl Acad Sci USA* 2011, **108:**4516-4522.
495 23. Rognes T, Flouri T, Nichols B, Quince C, Mahé F: **VSEARCH: a versatile open**
496 **source tool for metagenomics.** *Peerj* 2016, **4:**e2584.
497 24. Edgar RC: **Search and clustering orders of magnitude faster than BLAST.**
498 *Bioinformatics* 2010, **26:**2460-2461.
499 25. Edgar RC, Flyvbjerg H: **Error filtering, pair assembly and error correction for**
500 **next-generation sequencing reads.** *Bioinformatics* 2015, **31:**3476-3482.
501 26. Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA,
502 Clemente JC, Burkepile DE, Vega Thurber RL, Knight R, et al: **Predictive functional**
503 **profiling of microbial communities using 16S rRNA marker gene sequences.** *Nat*
504 *Biotechnol* 2013, **31:**814-821.
505 27. Knight R, Vrbanac A, Taylor BC, Aksenov A, Callewaert C, Debelius J, Gonzalez
506 A, Kosciolek T, McCall L, McDonald D, et al: **Best practices for analysing**
507 **microbiomes.** *Nat rev Microbiol* 2018, **16:**410-422.
508 28. Takebayashi A, Kimura F, Kishi Y, Ishida M, Takahashi A, Yamanaka A,
509 Takahashi K, Suginami H, Murakami T: **The association between endometriosis and**
510 **chronic endometritis.** *Plos One* 2014, **9:**e88354.
511 29. Baker JM, Chase DM, Herbst-Kralovetz MM: **Uterine Microbiota: Residents,**
512 **Tourists, or Invaders?** *Front Immunol* 2018, **9**.
513 30. Leyendecker G, Bilgicyildirim A, Inacker M, Stalf T, Huppert P, Mall G, Böttcher
514 B, Wildt L: **Adenomyosis and endometriosis. Re-visiting their association and**

 doi:10.20944/preprints201911.0147.v1

further insights into the mechanisms of auto-traumatisation. An MRI study. *Arch Gynecol Obstet* 2015, **291:**917-932.

31. Polatti F: **Bacterial vaginosis, Atopobium vaginae and nifuratel.** *Cur clini pharmacol* 2012, **7:**36.

32. Mendes-Soares H, Krishnan V, Settles ML, Ravel J, Brown CJ, Forney LJ: **Fine-scale analysis of 16S rRNA sequences reveals a high level of taxonomic diversity among vaginal Atopobium spp.** *Pathog Dis* 2015, **73**.

33. Walther-António MRS, Chen J, Multinu F, Hokenstad A, Distad TJ, Cheek EH, Keeney GL, Creedon DJ, Nelson H, Mariani A, Chia N: **Potential contribution of the uterine microbiome in the development of endometrial cancer.** *Genome Med* 2016, **8**.

34. Mu F, Harris HR, Rich-Edwards JW, Hankinson SE, Rimm EB, Spiegelman D, Missmer SA: **A Prospective Study of Inflammatory Markers and Risk of Endometriosis.** *Am J Epidemiol* 2018, **187:**515-522.

35. Chang CY, Lai MT, Chen Y, Yang CW, Chang HW, Lu CC, Chen CM, Chan C, Chung C, Tseng CC, et al: **Up-regulation of ribosome biogenesis by MIR196A2 genetic variation promotes endometriosis development and progression.** *Oncotarget* 2016, **7:**76713-76725.

36. Bhagirath AY, Li Y, Patidar R, Yerex K, Ma X, Kumar A, Duan K: **Two Component Regulatory Systems and Antibiotic Resistance in Gram-Negative Pathogens.** *Int J Mol Sci* 2019, **20:**1781.

37. Yang H, Kang K, Cheng C, Mamillapalli R, Taylor HS: **Integrative Analysis**

Reveals Regulatory Programs in Endometriosis. *Reprod Sci* 2014, **22:**1060-1072.

38. Yotova I, Hsu E, Do C, Gaba A, Sczabolcs M, Dekan S, Kenner L, Wenzl R, Tycko B: **Epigenetic Alterations Affecting Transcription Factors and Signaling Pathways in Stromal Cells of Endometriosis.** *Plos One* 2017, **12:**e170859.

39. Turner RD, Vollmer W, Foster SJ: **Different walls for rods and balls: the diversity of peptidoglycan.** *Mol Microbiol* 2014, **91:**862-874.

Figures captions

Fig 1:

PCA reduces complicated multidimensional distance into three-dimensional figures of OTUs sample distance in the largest extent but it may still hide some information. Because the three-dimensional figure is not able to fully explain multidimensional information, Y-axis (PC2), X-axis (PC1) and Z-axis (PC3) provides us the degree of explanation by percentage. These three PCA plots were conducted in three different levels. Green dots represent the samples from posterior fornix of cervix and red dots represent samples from cervical canals.

PCA: Principal Component Analysis

PC: Principal Component

Fig 2:

Four types of alpha diversity were analyzed including Chao1 index, Richness index,

 doi:10.20944/preprints201911.0147.v1

559 Shannon_e index and Simpson index. For each alpha index, four groups were analyzed
560 and $p<0.05$ was identified as statistically significant. The result showed no significance
561 in all groups.
562
563 Fig 3:
564 Each column was integrated by disease groups after the normalization, Order, Family
565 and Genus levels were analyzed and shown in the figure. Top 10 (abundance) OTUs in
566 all groups had been selected, and *Lactobacillus* was dominant in all groups.
567 *Lactobacillus* was less and *Atopobium* was more in adenmyosis-endometriosis than any
568 other groups.
569
570 Fig 4:
571 With OTU abundance >1% selection, 6 OTUs were displayed in this figure. Upper
572 half part of the circle shows four disease groups and the relative proportion of 6 OTUs.
573 Lower half part of the circle shows the proportion taken by each disease group in each
574 OTU. It should be noticed that the *Atopobium* in lower right, most of the *Atopobium*
575 was occupied by blue (AMEM group). This was studied in genus level.
576
577 Fig 5 and 6:
578 PCoA reduces complicated multiple-dimensional distance into two-dimensional
579 presentations of OTUs distance in the largest extent. Because the two-dimensional
580 figure is not able to fully explain multi-dimensional information, Y-axis and X-axis

 doi:10.20944/preprints201911.0147.v1

provides the degree of explanation by percentage. Weighted and Unweighted-Unifrac distances are qualitive matrixes and they only consider the present or absence of feature. PCoA: Principal co-ordinates analysis

Fig 7:

These figures show significant different microbiota by two-two comparison with $p<0.05$, mean proportion and 95% confidential interval were displayed in the figure. (a) This column displayed five microbiotas through two-two comparison among each group in genus level. (b) This column displayed two microbiotas through two-two comparison among each group in family level.

Supplementary data 1:

Supplementary data 2:

Alpha rarefaction is the assessment of the depth of sample size. The curve of all four groups were rising in the beginning of the curve and became horizonal in the end of curve, it means that the richness was still maximum and it will not rise if we enlarge the sample size further. Vertical segments are 95% confidential interval of each point. This alpha-rarefaction curve is aimed to estimate whether sample size is big enough to conduct alpha diversity analysis.

Supplementary fig 1:

PICRUSt is used to predict the functions through KO3 based on metagenome. The

 doi:10.20944/preprints201911.0147.v1

603 differences of function were analyzed in two-two comparison. Some distinctive
604 functions may take a small proportion that were hardly recognized. This identified
605 functions were filtered with $p<0.05$.
606 PICRUSt: Phylogenetic Investigation of Communities by Reconstruction of
607 Unobserved States
608 KO3: KEGG Orthology level 3
609
610