Distinction between vaginal and cervical microbiota in high-risk human papilloma virus infected women in China

Zhan Zhang
Capital Medical University Beijing Obstetrics and Gynecology Hospital
Ting Li
Capital Medical University Beijing Obstetrics and Gynecology Hospital
Dai Zhang
Peking University First Hospital
Xiaonan Zong
Capital Medical University Beijing Obstetrics and Gynecology Hospital
Huihui Bai
Capital Medical University Beijing Obstetrics and Gynecology Hospital
Hui Bi
Peking University First Hospital
Zhaohui Liu (17301255426@163.com)
Beijing Obstetrics and Gynecology Hospital, Capital Medical Univsersity https://orcid.org/0000-0001-7402-7095

Research article

Keywords: high-risk human papilloma virus, vaginal microbiota , cervical microbiota

Posted Date: January 11th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-60202/v2

Version of Record: A version of this preprint was published on March 25th, 2021. See the published version at https://doi.org/10.1186/s12866-021-02152-y.

Abstract

Background: High-risk human papilloma virus (hrHPV) is regarded as the main causal factor of cervical precancer and cancer when persistent infection is left untreated. Previous researches have confirmed that the vaginal microbiota was associated with HPV infection and the development of cervical lesions. A very recent study has revealed the microbiota at different parts of the female genital tract was closely related but different from each other. To analyze the distinction between vaginal and cervical microbiota of hrHPV(+) women in China, one hundred subjects were recruited including 20 healthy controls without HPV infection, 32 with other hrHPV(+), 38 with HPV16/18(+) and 10 with cervical carcinoma. Vaginal and cervical microbiota were separately tested through next-generation sequencing technology (NGS) targeting the variable region (V3-V4) of bacterial ribosome 16S rRNA gene.

Results: HrHPV(+) subjects tend to be accustomed to vaginal douching (p =0.001), show more frequent usage of sanitary pads (p =0.007), have more sex partners (p =0.047), be more sexually active (p =0.025) and more diversed in ways of contraception (p =0.001). Alpha diversity of cervical microbiota was higher than that of vagina. Cervical microbiota consisted a lower percentage of *Firmicutes* and a higher percentage of *Proteobacteria* compared to the vagina at the phylm level. *Sphingomonas* of α-*Proteobacteria* was almost below detection limit in the vagina whereas accounted for five to ten percent at hrHPV(-) cervix (P<0.001) and reversely associated with hrHPV infection (P<0.05). *Pseudomonas* of γ-*Proteobacteria* could hardly be seen in the normal vagina and shared a little bit percentage in the normal cervix, but significantly higher in the HPV16/18 infected (P<0.001) and cancerous cervix (P<0.05). BV associated anaerobes like *Gardnerella*, *Prevotella*, *Atopobium* and *Sneathia* did not exhibit cervical specificity.

Conclusions: Cervical microbiota has its uniqueness from that of vagina in bacterial communities presenting a higher proportion of *Proteobacteria*, of which *Pseudomonas* is positively while *Sphingomonas* was negatively associated with hrHPV infection and cervical cancer. It is of great improtance to deeply explore the cervical microbiota of hrHPV infected women.

Background

Human papilloma virus (HPV) is a double-stranded DNA virus that only infects the human body, and two types have been identified, namely, the skin type and mucosal type, comprising over 100 subtypes [1]. More than 40 mucosal types of HPV can infringe on the human reproductive system, and 15 high-risk HPV (hrHPV) types have been demonstrated to be related to the cervical lesion: HPVs 16, 18, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66 and 68 [1]. Approximately 50% and 20% of cervical cancers are induced by HPV 16 and 18 respectively, and these two subtypes are considered to be extremely high-risk types [2]. Persistent HPV infection plays a pivotal role in cervical cancer development. The progression from hrHPV infection to squamous intraepithelial lesion (SIL) and cancer, can generally last years or even decades [3], and within the time window, the preventative or therapeutic intervention could be executed.

As the world's largest developing country, about 300 million women in China needs cervical cancer screening every year. Primary prevention strategy represented by HPV vaccine in China is still in its infancy [4]. China has not incorporated the HPV vaccine in the National Immunization Program. After the United States licensure in 2006, the first HPV vaccine was approved in China until 2016. Now information about vaccine and its acceptance among Chinese women is scarce [4]. Current situation in primary prevention of hrHPV is still not optimistic. So we should strengthen the efforts of secondary and tertiary prevention, and at the same time look for deeper factors of hrHPV infection and cervical lesion, which will help to provide theoretical basis for innovation in the prevention and control of hrHPV and cervical lesion.

In addition to HPV infection, several other related factors were also involved in the cancerous progression. Studies [5,6] have paid attention to the following factors and confirmed their association with cervical lesion: social economical factors, hygienic habits, sexual and parity etc. With the development of microbiological detection technology especially new generation sequencing (NGS), there has been increasing concerns in recent years that genital microbial environment maybe associated with HPV infection and cervical lesions [7].

It has been recently proposed that abnormal vaginal microbiota played a significant role in the development of cervical neoplasm. In the female lower reproductive tract, health is more commonly associated with low microbial diversity and dominance by only one or a few species of *Lactobacillus* [8]. Lee [9] analyzed the relationship between HPV infection and vaginal flora for the first time and discovered a higher diversity of vaginal microbiota in HPV infected women. Soon after, Brotman [10] found that the type of vaginal microbiota might be associated with HPV clearance or persistence. Thus, lots of experts have begun to pay attention to microdysbiosis of the female lower genital tract and infer that vaginal microbiota disturbance might directly relate to HPV acquisition and even to cervical cancer. A very recent study [11] has revealed that microbiota at different parts of the female genital tract might be closely related whereas distinct from each other, changing from vagina to cervix, endometrium, fallopian tubes and peritoneal fluid. Most of the studies to date preferred to refer the sample as “cervicovaginal” instead of discussing “cervical” and “vaginal” samples respectively. And no such research has been done on the distinction between cervical and vaginal microbiota of hrHPV infected women in Chinese population, which has led us to design this project to explore the continuum and distinction between the cervical and vaginal microbiota in hrHPV infected women in China. This research would be meaningful for further researches to explore the role of microbiota in cervical carcinogenesis [12].

Results

Demographics

In order to characterize the cervical and vaginal microbiota in hrHPV(+) Chinese women, we obtained vaginal and cervical samples from 100 subjects and divided them into four groups, namely, the **n**ormal/control group (Group N, n =20), the **o**ther hrHPV group (Group O, n =32), **H**PV 16/18 group (Group H, n =38) and **c**ervical cancer group (Group C, n =10). HPV 16/18 and other high-risk subtypes were

specifically separated since HPV 16/18 are extremely high-risk subtypes and cause nearly seventy percent of cervical carcinoma [2]. Subjects of each group were age matched (Tab.1, $P = 0.289$). All enrolled subjects underwent regularly TCT and HPV tests (Tab.1). Apart from HPV16/18 the most frequent HPV subtypes were HPV 52 (37.5%), 58 (15.6%), 33 (12.5%) and 53 (12.5%). Prevalence rate of HPV 16 and 18 were 81.6% and 47.4% respectively, in which some were coinfected with both. Subjects of group C were diagnosed with cervical squamous carcinoma staged from Federation of International Gynecology and Obstetrics (FIGO) Ia1 to IIa2 (Tab.1). Subjects of group H and C tended to have more frequent usage of sanitary pads (P =0.007), more sex partners (P =0.047), more frequent intercourse (P =0.025) and be more accustomed to vaginal douching (P =0.001). People infected with hrHPV had a higher proportion of vaginitis history (P =0.002). Condom use was significantly lower among hrHPV(+) individuals and contraceptive methods were more varied (P =0.001).

Diversity of cervical microbiota is different from that of vagina

Microbiota sequencing was performed targeting the V3-V4 region of the 16S rRNA genes using Illumina Miseq platform. Herein cervical microbiota was separately discussed from vaginal microbiota. For convenience, vaginal microbiota of four groups were respectively abbreviated to Vn, Vo, Vh and Vc, and cervical microbiota correspondingly abbreviated to Cn, Co, Ch and Cc. Subscripts like “n, o, h and c” repectively represented the **n**ormal group, the **o**ther hrHPV(+) group, the **h**rHPV16/18 group and the **c**ancer group. We used Shannon index to represent the Alpha diversity of species. The higher the Shannon, the more diversed the microbe in the sample or group was. Results suggested the Shannon index of Cn was significantly higher than Vn, demonstrating a higher microbe diversity of healthy cervix than vagina (Fig. 1a, P =0.019). The same trends were observed in Co vs Vo (Fig. 1b, P =0.018) and Ch vs Vh (Fig. 1c, P =0.034). Whereas no significant difference has been found between Cc and Vc (Fig. 1d, P =0.466). To clarify whether cervical microbiota is different from vaginal microbiota, Beta diversity analysis was also performed. In this part the Unifrac distance was calculated to estimate the evolutionary differences of species between different groups. The distances boxplot showed the significant diffierence between cervical and vaginal microbiota (Fig.2 Vn vs Cn, P 0.001; Vo vs Co, P 0.001; Vh vs Ch, P 0.001; Vc vs Cc, P 0.05), which revealed vaginal and cervical microbes consisted of different communities from an evolutionary perspective.

Proteobacteria are much more prevalent at cervix compared to vagina

Six major taxa at the phylm level were observed in vaginal and cervical microbiota, namely, *Firmicutes*, *Actinobacteria*, *Bacteroidetes*, *Fusobacteria*, *Proteobacteria* and *Tenericutes* (Fig. 3). Linear discriminant analysis Effect Size (LEfSe) analysis showed that cervical microbiota consisted a lower percentage of *Firmicutes and* a much higher percentage of *Proteobacteria* compared to the vagina in HPV(-) subjects (Fig. 4a), which indicates the *Proteobacteria* as a special phylm at normal cervix. *Proteobacteria* are a large family of bacteria. To explore the particularity of *Proteobacteria* at hrHPV(+) cervix, we further conducted LEfSe analysis and observed that γ-*Proteobacteria* were more prevelent at cancerous cervix (Fig. 4b).

In order to find the target genus reside in the cervix that associated with hrHPV infection and cervical cancer, we systematically compared the presence of some representative bacteria in the vagina and cervix of hrHPV(+) subjects, like *Lactobacillus*, *Sphingomonas* of α-*Proteobacteria*, *Pseudomonas* of γ-*Proteobacteria* and bacterial vaginosis (BV) related anaerobes as *Gardnerella*, *Prevotella*, *Atopobium* and *Sneathia*. We have observed reduced *Lactobacillus* was associated with cervical cancer (Fig. 5a, Vn vs Vc, $P \leq 0.05$; Cn vs Cc, $P \leq 0.05$; Vo vs Vc, $P \leq 0.01$; Co vs Cc, $P \leq 0.01$; Vh vs Vc, $P \leq 0.05$; Ch vs Cc, $P \leq 0.05$). A lower level of *Lactobacillus* was seen at cervix than the vagina in both hrHPV(+) and hrHPV(-) subjects (Fig. 5b, Vn vs Cn, $P \leq 0.01$; Vo vs Co, $P \leq 0.001$; Vh vs Ch, $P \leq 0.001$).

Sphingomonas and *Pseudomonas* were selected as representative genus in *Proteobacteria* because we found their association with hrHPV infection and cervical cancer. *Sphingomonas* of α-*Proteobacteria*, which was almost below detection limit in the vagina (Fig. 5c) whereas accounted for five to ten percent at hrHPV(-) cervix (Fig. 5d, Vn vs Cn, $P \leq 0.001$) and reversely associated with hrHPV infection (Fig. 5c, Cn vs Co/Ch/Cc, $P \leq 0.05$). *Pseudomonas* could hardly be seen in the normal vagina (Fig. 5e) and shared a little bit percentage in the HPV16/18(+) and cancerous vagina (Fig. 5e). Whereas *Pseudomonas* was relatively high in the HPV16/18(+) and cancerous cervix (Fig. 5e, Cn vs Ch, $P \leq 0.001$; Co vs Ch, $P \leq 0.001$; Cn vs Cc, $P \leq 0.05$; Co vs Cc, $P \leq 0.01$. Fig. 5f, Vh vs Ch, $P \leq 0.001$; Vc vs Cc, $P \leq 0.001$).

BV related anaerobes showed similar changes in both vaginal and cervical microbiota of hrHPV(+) subjects and did not exhibit “cervical specificity” (Fig. 6 b, d, f, h). *Prevotella* was higher mainly in the cancerous vagina and cervix (Fig. 6a, Vo vs Vc, $P \leq 0.05$; Cn vs Cc, $P \leq 0.05$). *Gardnerella* shared a higher percentage in HPV16/18(+) cervix and cancerous vagina/cervix(Fig. 6c, Vn vs Vc, $P \leq 0.05$; Cn vs Cc, $P \leq 0.01$; Cn vs Ch, $P \leq 0.05$). *Atopobium* was higher in HPV16/18(+) and cancerous vagina/cervix (Fig. 6e, Vn vs Vh, $P \leq 0.05$; Vn vs Vc, $P \leq 0.01$; Cn vs Ch, $P \leq 0.01$; Cn vs Cc, $P \leq 0.01$), and more prevalent in other hrHPV infected cervix (Fig. 6e, Cn vs Co, $P \leq 0.01$). While *Sneathia* were significantly higher in all hrHPV infected vagina/cervix regardless of HPV subtypes (Fig. 6g, Vn vs Vo, $P \leq 0.05$; Vn vs Vh, $P \leq 0.05$; Cn vs Co, $P \leq 0.001$; Cn vs Ch, $P \leq 0.05$; Cn vs Cc, $P \leq 0.01$).

Function of cervical microbiota is more active than that of vagina

Sequences acquired by NGS were translated to corresponding functions using software R by paralleling to Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. We observed that the microbial functions at cervix were more complicated and active than those in vagina. Some special functions involved the amino acid metabolism, carbohydrate metabolism, membrane transport, replication and repair, and gene information processing were observed relatively vigorous at cervix (Fig. 7). These functions are very important for virus replication, integration and development of cervical lesions.

Discussion

With the rapid development of microbial detection methods especially the application of NGS, we are increasingly aware of the importance of microorganism to human health [13-14]. Certain members of bacteria in the lower genital tract are believed to be benificial for women against infection and

pathogenesis. Healthy vagina is more commonly associated with low microbial diversity and dominated by one or a few species of *Lactobacillus* [15-17]. The primary defense mechanisms of the lower genital mucosa are antimicrobial peptides, a pH of less than 4.5, and a microbial community dominated by *Lactobacillius*. An imbalance in these defenses could result in pathological alterations of lower genital tract [7].

HPV is a unique health concern because it is one of very few infections that can definitely lead to cancer. In most cases the immune system clears the virus on its own within 6-18 months [18-19]. It takes a long time from HPV infection to the development of cancer, which gives us oppurtunity to prevent this deterioration process. Over the past decade, evidences have suggested that the vaginal microbiome also play a role in cervical carcinogenesis [20]. Emerging studies have revealed associations between the non-*Lactobacillus* dominant vaginal microbiota and HPV infection and persistence [21]. Whereas previous studies on lower genital microbiota only focused on vaginal microbiota, or did not separate the cervical and vaginal flora, simply called it “cervicovaginal microbiota” [22-24], which may be because it is generally believed there is no difference between the vaginal and cervical microbiota and that contamination is difficult to avoid when separately sampling.

The highlight of our study is that this is the first study that has discussed the distinction between cervical and vaginal microbiota of hrHPV(+) women in China. We are enlightened by the study of Peking University Shenzhen Hospital in 2017 [11], which broke the idea that the upper genital tract was sterile and revealed differences in the microbiota of different parts of female reproductive tract. They systematically sampled the discharge at six sites within the female genital tract from a large cohort of Chinese women of reproductive age. The six part are respectively the lower third of vagina, the posterior fornix, the cervical canal, the endometrium, the left/right fallopian tubes and the peritoneal fluid from the pouch of Douglas. They have clarified that at the phylum level the *Firmicutes* dominated lower reproductive tract contrasted the large proportions of *Proteobacteria*, *Actinobacteria* and *Bacteroidetes* in the upper reproductive tract. We have adopted the sampling methods in this study and cervical discharges are strictly taken from the cervical canal to avoid contamination. Similarly our research revealed cervical microbiota consisted a lower percentage of *Firmicutes and* a much higher percentage of *Proteobacteria* compared to the vagina. *Proteobacteria* account for approximately one percent of normal vaginal microbiota but more than ten percent of normal cervical microbiota, demonstrating itself to be a particular phylum at cervix. We speculate that lower percentage of *Firmicutes-Lactobacillius* affects the production of lactic acid and hydrogen peroxide [25], thus the pH of the cervix tends to be higher than the vagina, which further changes the composition of other strains.

Since *Proteobacteria* was a special phylm in the cervix than vagina. Are particular genuses of *Proteobacteria* associated with hrHPV infection or cervical cancer? There have been no such studies on cervical *Proteobacteria* and hrHPV infection. *Proteobacteria* [26] comprises quite a large communities and can be divided into α-, β- and γ-*Proteobacteria*. We have noticed *Pseudomonas* was more prevalent in the hrHPV(+) especially in the cancerous cervix. *Pseudomonas* belongs to the γ-*Proteobacteria*, and has long been believed to be an opportunistic pathogen of human urogenital system. Jeff [27] has studied the

role of infectious factors in cervical cancer using *Pseudomonas aeruginosa* as the bacterial tool and *Lactobacillus* as control. They observed upregulated expression of integrins in cervical cancer tissues and found *Pseudomonas aeruginosa* could promote the expression of integrins in cervical cancer cell lines, while the control group of *Lactobacillus* showed no change. This result indicates the potential role of *Pseudomonas* in promoting the development of cervical lesions. And this is the only basic experimental research on Pseudomonas' pathogenesis on cervical lesions. Of course we need more researches to further expand the mechanism in detail.

Another genus of α-*Proteobacteria*, *Sphingomonas* was more frequent at normal cervix and infrequent at hrHPV(+) and cancerous cervix. Studies on *Sphingomonas* that we have seen are mainly distributed in the field of environmental science. We have not yet seen any studies on *Sphingomonas* in the human reproductive system. *Sphingomonas* are often isolated from petroleum contaminated soils due to their unique abilities to degrade polycyclic aromatic hydrocarbons (PAHs) [28], which has been recognized as a definite carcinogen ~~in the world~~. While another study [29] suggests that both HPV infection and PAHs are critical factors in the development of cervical cancer. PAHs have potential to ~~cooperate~~ coordinate with HPV to aggravate carcinogenesis at all stages. Therefore, it is not clear whether *Sphingomonas* could play a protective role in the development of cervical lesions by degrading PAHs, which is worthy exploring in further study. If this hypothesis becomes a reality, we may be able to prevent HPV infection and cervical lesions from another perspective.

Compared to *Proteobacteria*, BV related anaerobes showed similar changes in the hrHPV(+) vagina and cervix, and have not showed cervical particularity. *Sneathia* have the most sensitive association with hrHPV infection, regardless of HPV subtypes. Actually *Sneathia* is the first genus to be identified as a target bacteria for HPV infection early in 2013 [9]. Intriguingly, *Sneathia* is the only bacteria that is enriched in genital tract throughout the process of cervical carcinogenesis [30]. By comparison, *Gardnerella*, *Prevotella* and *Atopobium* are more prevalent mainly in HPV16/18(+) or cancerous vagina/cervix. This may explain why HPV16/18 are more carcinogenic, as people with HPV16/18(+) have a more disturbed microenvironment. While we need more prospective studies to confirm it. By comparing BV associated anaerobes with certain strains of *Proteobacteria* like *Sphingomonas* and *Pseudomonas*, we reconfirmed that there existed community differences between the cervical and vaginal microbiota. It is of great significance to deeply analyze the difference between vaginal and cervical microbiota, to find the cervical target genus and to explore their functional mechanism so as to better prevent HPV infection.

Methods

Study cohort and sample collection

The study was conducted in accordance with the Declaration of Helsinki and its current amendments and the protocol was approved by the medical ethics committee of Peking University First Hospital. All subjects provided written informed consent and there was no financial compensation. One hundred women of reproductive age in the Gynecological clinic were recruited and divided into four groups

according to routine cervical cancer screening results. The **n**ormal group (Group N) comprised 20 women whose Thinprep cytologic test (TCT) and HPV were both negative. HPV16/18 group (Group H) comprised 38 **h**rHPV16/18 (+) women and colposcope biopsies showed no cancerous lesions. **O**ther hrHPV group (Group O) comprised 32 women of hrHPV except for HPV 16/18 and biopsies showed no cancerous lesions. **C**ancer group (Group C) comprised 10 women of cervical carcinoma.

Inclusion criteria: women of reproductive age; having sexual experience; having regular menstruation; mid-follicular phase; no usage of any medications within one week; no vaginal douching, cervical treatment or sexual intercourse within 72 hours. Exclusion criteria: women during pregnancy, lactation or menopause and women with chronic diseases that need long-term medication.

All participants were required to fill in questionnaires, including age, educational level, occupation, economic condition, hygiene practices, sexual activity, history of vaginitis and cervical cancer screening results. Discharge of vagina and cervix were respectively collected and reserved with separate sterile cotton swabs and eppendorf (EP) tubes containing normal saline, stored at -80°C and transported on dry ice to Sangon Biotech-Shanghai for NGS. Vaginal discharge was obtained from up one-third of vagina. Cervical discharges are strictly taken from the cervical canal [11]. To avoid contamination, we wiped the surface secretions of cervix with a sterile cotton swab before formal sampling. Another swab~~s were~~ was used for smearing, Gram staining and oil lens observation to evaluate the vaginal microecology and numbers of cervical leukocytes. The Nugent score was adopted to diagnose BV (Nugent score 7-10: BV; 4-6: BV intermediate; 1-3: normal). Vulvovaginal candidiasis (VVC) was indicated when hypha or spores were discovered, and Trichomonas vaginitis (TV) was indicated when Trichomonas was seen under oil lens. Cervicitis was indicated when the average number of leukocytes was more than 10/high power field. For the convenience of understanding, the vaginal microbiota of four groups (Group N O H C) were abbreviated to Vn, Vo, Vh, Vc, and the cervical microbiota of four groups are abbreviated to Cn, Co, Ch, Cc.

DNA extraction and 16S rRNA V3-V4 gene sequencing

DNA extraction was performed according to the instructions of the OMEGA E.Z.N.ATM Mag-Bind Soil DNA Kit. DNA integrity was detected by agarose gel electrophoresis. V3-V4 region of the 16S rRNA genes were amplified by polymerase chain reaction (PCR) with an universal forward primer and unique barcode primer [31] (V3-341F : CCCTACACGACGCTCTTCCGATCTG (barcode) CCTACGGGNGGCWGCAG; V4-805R : GACTGGAGTTCCTTGGCACCCGAGAATTCCA (barcode) GACTACHVGGGTATCTAATCC).

The first amplification was performed under the following conditions: 3 min of denaturation at 94 °C; 5 cycles of denaturation at 94°C for 30s, annealing at 45°C for 20s, and elongation at 65°C for 30s; 20 cycles of denaturation at 94°C for 20s, annealing at 55°C for 20s, and elongation at 72°C for 30s; and a final extension at 72°C for 10min. Illumina bridge PCR compatible primers were introduced in the second amplification as follows: 3min of denaturation at 95°C; 5cycles of denaturation at 94°C for 20s, annealing at 55°C for 20s, and elongation at 72°C for 30s; and a final extension at 72°C for 10min. Amplicons were purified using the AMPure XP beads and DNA quantitation was performed using the

Qubit 3.0 DNA Kit, 10 ng of DNA extracted from each sample was sequenced using the Illumina Miseq 2×300 bp platform.

Raw data were processed as follows: remove joint sequences of primers, splicing sequences according to the overlap, identify sample data by barcode and remove chimeras and nonspecific sequences to achieve quality control. Operational Taxonomic Unit (OTU) clustering was performed at 97% similarity level. Software used: Cutadapt, PEAR [32], Prinseq [33], Usearch [34] and Uchime [35].

Bioinformatics analysis

Microorganism taxonomy: Species taxonomy was performed based on the Ribosomal Database Project (RDP) classifier [36] and Bergey's taxonomy. Microbes were analyzed at six levels, namely, domain, phylum, class, order, family and genus. Taxonomy diagrams were drawn using software R [37].

Alpha and beta diversity: Alpha diversity was used to evaluate the diversity of species within each sample or each group. Calculated shannon index of each group. The greater the shannon value was, the higher the diversity of community would be. Beta diversity was used to measure the evolutionary distances between different samples or groups evaluated by UniFrac distance. Distance of UniFrac is between 0 and 1, and the higher the value, the further the evolutional distance. Software used: Muscle [38], FastTree [39], mothur [40] and R [37].

Linear discriminant analysis Effect Size (LEfSe) analysis: LEfSe analysis could best explain the community difference in each group. Statistical methods: Kruskal-Wallis rank-sum test, (unpaired) Wilcoxon rank-sum test and Linear Discriminant Analysis. Software used: LEfSe [41].

Functional analysis: Composition of functional genes were calculated by KEGG [42] database. Software used: PICRUSt [43].

Statistical analysis of demographics data

SPSS 19.0 and Graphpad Prism 5 were used for data analysis. Age was analyzed by ANOVA, the other items were analyzed by Chi-square test or Fisher exact probability test. For the analysis of genus differences, the Mann Whitney test and Wilcoxon Signed Rank test were respectively used for the comparative analysis within and between the vaginal and cervical microbiota. $P < 0.05$ was statistically significant.

Conclusions

This is the first study that has paid special attention to cervical microbiota of hrHPV(+) Chinese women, and distinguish it from the vaginal microbiota. The results revealed *Proteobacteria* to be a particular phylm at cervix than the vagina. And *Sphingomonas* of α-*Proteobacteria* has potential to play a protective role ~~from~~ during hrHPV infection, while *Pseudomonas* of γ-*Proteobacteria* is positively associated with hrHPV infection and cervical cancer. These findings will provide new ideas for the prevention of hrHPV

from the perspective of microecology. This project also has some shortcomings, such as the limitations of cross-sectional study and small sample size etc. Large-scale prospective clinical trials need to be implemented in future to discover the changes of microbiota longitudinally in the chronic process of persistent hrHPV infection and to explore the predictive and therapeutic value of specific genera on hrHPV infection and cervical lesions. This is a microbiological age and microecological prevention and therapy will become possible.

Declarations

Ethics approval and consent to participate

The study has been approved by the medical ethics committee of Peking University First Hospital. All participants have provided written informed consent.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This research has no financial support.

Authors' contributions

ZZ participated in study design, literature search, ethical application, patients recruitment, specimen collection, data analysis and manuscript writing. TL analyzed and interpreted the patient data and participated in manuscript writing. DZ participated in study design and patients recruitment. XZ and HB (Huihui Bai) provided help in literature search, patients recruitment and specimen collection. HB (Hui Bi) participated in study design, patients recruitment, specimen collection and manuscript review. ZL participated in study design, patients recruitment, specimen collection, data analysis, manuscript editing and manuscript review. All authors have read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Schiffman M,Castle PE,Jeronimo J, Rodriguez AC, Wacholder S.Human papillomavirus and cervical cancer. Lancet 2007; 370:890-907.

2. de Sanjose S,Quint WG,Alemany L, Geraets DT, Klaustermeier JE, Lloveras B,et al.Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study.Lancet Oncol 2010;11: 1048-56.

3. Kyrgiou M, Mitra A, Moscicki AB. Does the vaginal microbiota play a role in the development of cervical cancer?. Transl Res. 2017;179:168-82.

4. You D, Han L, Li L, Hu J, Zimet GD, Alias H,et al.Human Papillomavirus (HPV) Vaccine Uptake and the Willingness to Receive the HPV Vaccination among Female College Students in China: A Multicenter Study.Vaccines (Basel) 2020;8.pii: E31.

5. Yang J, Wang W, Wang Z, Wang Z, Wang Y, Wang J,et al.Prevalence, genotype distribution and risk factors of cervical HPV infection in Yangqu, China: a population-based survey of 10086 women.Hum Vaccin Immunother 2019;6:1-8.

6. Ciccarese G, Herzum A, Rebora A, Drago F.Prevalence of genital, oral, and anal HPV infection among STI patients in Italy.J Med Virol 2017;89:1121-24.

7. Audirac-Chalifour A, Torres-Poveda K, Bahena-Román M, Téllez-Sosa J, Martínez-Barnetche J, Cortina-Ceballos B, et al.Cervical Microbiome and Cytokine Profile at Various Stages of Cervical Cancer: A Pilot Study.PLoS One 2016; 11:e0153274.

8. Mitra A, MacIntyre DA, Marchesi JR, Lee YS, Bennett PR, Kyrgiou M.The vaginal microbiota, human papillomavirus infection and cervical intraepithelial neoplasia: what do we know and where are we going next?Microbiome

2016;4:58.

9. Lee JE,Lee S,Lee H, Song YM, Lee K, Han MJ, et al.Association of the vaginal microbiota with human papillomavirus infection in a Korean twin cohort. PLoS One 2013; 8:e63514.

10. Brotman RM,Shardell MD,Gajer P, Tracy JK, Zenilman JM, Ravel J, et al. Interplay between the temporal dynamics of the vaginal microbiota and human papillomavirus detection.J Infect Dis 2014;210:1723-33.

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

12. Adebamowo SN, Dareng EO, Famooto AO, Offiong R, Olaniyan O, Obende K,et al.Cohort Profile: African Collaborative Center for Microbiome and Genomics Research's (ACCME's) Human Papillomavirus (HPV) and Cervical Cancer Study.Int J Epidemiol 2017;46:1745-1745j.

13. Cong J, Zhang X. How human microbiome talks to health and disease. Eur J Clin Microbiol Infect Dis.2018;37(9):1595-601.

14. Integrative HMP (iHMP) Research Network Consortium. The Integrative Human Microbiome Project. Nature. 2019; 569(7758): 641-8.

15. Witkin SS, Linhares IM.Why do lactobacilli dominate the human vaginal microbiota? BJOG 2017;124: 606-11.

16. Godha K, Tucker KM, Biehl C,Archer DF, Mirkin S.Human vaginal pH and microbiota: an update.Gynecol Endocrinol 2018; 34:451-5.

17. Kyrgiou M, Mitra A, Moscicki AB.Does the vaginal microbiota play a role in the development of cervical cancer?Transl Res 2017;179:168-82.

18. Jin J.HPV Infection and Cancer.JAMA 2018;319:1058.
19. Trottier H,Ferreira S,Thomann P, Costa MC, Sobrinho JS, Prado JC,et al.Human papillomavirus infection and reinfection in adult women: the role of sexual activity and natural immunity.Cancer Res 2010;70:8569-77.
20. Ilhan ZE, Łaniewski P, Thomas N, Roe DJ, Chase DM, Herbst-Kralovetz MM. Deciphering the complex interplay between microbiota, HPV, inflammation and cancer through cervicovaginal metabolic profiling. EBioMedicine. 2019;44:675-90.
21. Łaniewski P, Ilhan ZE, Herbst-Kralovetz MM. The microbiome and gynaecological cancer development, prevention and therapy. Nat Rev Urol. 2020;17(4):232-50.
22. Shannon B, Yi TJ, Perusini S, Gajer P, Ma B, Humphrys MS, Thomas-Pavanel J, Chieza L, Janakiram P, Saunders M, Tharao W, Huibner S, Shahabi K, Ravel J, Rebbapragada A, Kaul R. Association of HPV infection and clearance with cervicovaginal immunology and the vaginal microbiota. Mucosal Immunol. 2017;10(5):1310-19.
23. Usyk M, Zolnik CP, Castle PE, Porras C, Herrero R, Gradissimo A, Gonzalez P, Safaeian M, Schiffman M, Burk RD; Costa Rica HPV Vaccine Trial (CVT) Group. Cervicovaginal microbiome and natural history of HPV in a longitudinal study. PLoS Pathog. 2020;16(3):e1008376.
24. Shannon B, Gajer P, Yi TJ, Ma B, Humphrys MS, Thomas-Pavanel J, Chieza L, Janakiram P, Saunders M, Tharao W, Huibner S, Shahabi K, Ravel J, Kaul R. Distinct Effects of the Cervicovaginal Microbiota and Herpes Simplex Type 2 Infection on Female Genital Tract Immunology. J Infect Dis. 2017;215(9):

1366-75.

25. Witkin SS, Linhares IM. Why do lactobacilli dominate the human vaginal microbiota? BJOG. 2017;124(4):606-11.

26. Degli Esposti M.Bioenergetic evolution in proteobacteria and mitochondria. Genome Biol Evol 2014;6:3238-51.

27. Werner J,Decarlo CA,Escott N, Zehbe I, Ulanova M.Expression of integrins and Toll-like receptors in cervical cancer: effect of infectious agents.Innate Immun 2012;18:55-69.

28. Zhou L, Li H, Zhang Y, Han S, Xu H.Sphingomonas from petroleum-contaminated soils in Shenfu, China and their PAHs degradation abilities.Braz J Microbiol 2016;47:271-8.

29. Zhang C, Luo Y, Zhong R,Law PTY, Boon SS, Chen Z,et al. Role of polycyclic aromatic hydrocarbons as a co-factor in human papillomavirus-mediated carcinogenesis.BMC Cancer 2019;19:138.

30. Łaniewski P, Barnes D, Goulder A, Cui H, Roe DJ, Chase DM, Herbst -Kralovetz MM. Linking cervicovaginal immune signatures, HPV and microbiota composition in cervical carcinogenesis in non-Hispanic and Hispanic women. Sci Rep. 2018;8(1):7593.

31. Herlemann DP, Labrenz M, Jürgens K, Bertilsson S, Waniek JJ, Andersson AF. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. ISME J. 2011;5(10):1571-9.

32. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics. 2014;30(5):614-20.

33. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic

datasets. Bioinformatics. 2011;27(6):863-4.

34. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26(19):2460-1.

35. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics. 2011;27(16):2194-200.

36. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol. 2007;73(16):5261-7.

37. Chan BKC. Data Analysis Using R Programming. Adv Exp Med Biol. 2018;1082:47-122.

38. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research. 2004;32(5):1792-7.

39. Price MN, Dehal PS, Arkin AP. FastTree 2-Approximately Maximum-Likelihood Trees for Large Alignments. Poon AFY, ed. PLoS ONE. 2010;5(3):e9490.

40. Schloss PD, Westcott SL, Ryabin T, et al. Introducing mothur: Open-Source, PlatformIndependent, Community-Supported Software for Describing and Comparing Microbial Communities . Applied and Environmental Microbiology. 2009;75(23):7537-41.

41. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. Genome Biology. 2011;12(6):R60.

42. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017;45(D1):D353-61.

43. Langille MGI, Zaneveld J, Caporaso JG, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nature biotechnology. 2013;31(9):814-21.

List Of Abbreviations

HPV: Human papilloma virus; hrHPV: high-risk HPV; NGS: Next-generation sequencing technology; SIL: Squamous intraepithelial lesion; TCT: Thin-prep cytology test; Ascus: Atypical squamous cell of undetermined significance; LSIL: low-grade SIL; HSIL: high-grade SIL; ASC-H: Atypical squamous cell-cannot exclude HSIL; FIGO: Federation of International Gynecology and Obstetrics; BV: Bacterial vaginosis; KEGG: Kyoto Encyclopedia of Genes and Genomes; PAH: Polycyclic aromatic hydrocarbon; EP: Eppendorf; VVC: Vulvovaginal candidiasis; TV: Trichomonas vaginitis; PCR: Polymerase chain reaction; OTU: Operational Taxonomic Unit; RDP: Ribosomal Database Project; LEfSe: Linear discriminant analysis Effect Size.

Table

Table 1: Demographics of participants

	Normal (n=20)	Other hrHPV(+) (n=32)	HPV16/18(+) (n=38)	Cervical cancer (n=10)	*P**
Mean age (years)	38.35±3.72	36.75±6.15	35.73±6.89	38.80±3.08	0.289
TCT					
Normal	20	9（28.1%）	22（57.9%）	-†	
Ascus	-	11（34.3%）	3（7.9%）	-	
LSIL	-	8（25.0%）	4（10.5%）	-	
HSIL	-	2（6.3%）	7（18.4%）	-	
ASC-H	-	2（6.3%）	2（5.3%）	-	
HPV subtypes‡					
16	-	-	31（81.6%）	8(80.0%)	
18	-	-	8（47.4%）	3(30.0%)	
16and18	-	-	1(2.6%)	1(10.0%)	
31	-	1（3.1%）	1(2.6%)	0	
33	-	4（12.5%）	2(5.3%)	0	
35	-	1（3.1%）	0	1(10.0%)	
51	-	2（6.3%）	3(7.9%)	0	
52	-	12（37.5%）	3(7.9%)	0	
53	-	4（12.5%）	2(5.3%)	0	
56	-	3（9.4%）	2(5.3%)	0	
58	-	5（15.6%）	2(5.3%)	0	
59	-	3（9.4%）	0	0	
66	-	2（6.3%）	0	0	
68	-	3（9.4%）	0	0	
Vaginitis or not					0.002
Normal	16（80.0%）	27（84.4%）	25（65.8%）	3（30.0%）	
Bacterial vaginosis	2（10.0%）	5（15.6%）	13（34.2%）	5（50.0%）	
Abnormal flora	1（5.0%）	0	0	1（10.0%）	
Flora suppression	1（5.0%）	0	0	1（10.0%）	

Leukocytes at cervix					0.001
0-10	14（70.0%）	12（37.5%）	12（31.6%）	1（10.0%）	
＞10	6（30.0%）	20（62.5%）	26（68.4%）	9（90.0%）	
Cervical biopsy					0.129
Normal	20	0	0	-	
Cervicitis	-	14（43.8%）	11（28.9%）	-	
LSIL	-	11（34.4%）	10（26.3%）	-	
HSIL	-	7（21.8%）	17（44.8%）	-	
Cancer	-	0	10	10	
Educational level					0.478
＜Bachelor	9（45.0%）	12（37.5%）	17（44.7%）	5（50.0%）	
Bachelor	10（50.0%）	19（59.4%）	15（39.5%）	5（50.0%）	
≥Master	1（5.0%）	1（3.1%）	6（15.8%）	0	
Monthly income (¥)					0.295
＜5000	8（40.0%）	9（28.1%）	19（50.0%）	6（60.0%）	
5000-10000	10（50.0%）	17（53.1%）	12（31.6%）	4（40.0%）	
＞10000	2（10.0%）	6（18.8%）	7（18.4%）	0	
Occupation					0.838
Medical service	2（10.0%）	2（6.3%）	0	0	
Economics	2（10.0%）	5（15.6%）	5（13.2%）	0	
Education	1（5.0%）	2（6.3%）	4（10.5%）	1（10.0%）	
Art	1（5.0%）	1（3.1%）	1（2.6%）	0	
Worker/Farmer	14（70.0%）	22（68.7%）	28（73.7%）	9（90.0%）	
Frequency of cleaning vulva					0.306
＜1 time per week	0	4（12.5%）	3（7.9%）	0	
2-3 times per week	6（30.0%）	5（15.6%）	10（26.3%）	5（50.0%）	
Everyday	14（70.0%）	23（71.9%）	25（65.8%）	5（50.0%）	
Way of cleaning vulva					0.195
Clean water	19（95.0%）	23（71.9%）	31（81.6%）	9（90.0%）	
Lotion	1（5.0%）	9（28.1%）	7（18.4%）	1（10.0%）	

History of vaginal douching					0.001
Yes	0	11（34.4%）	19（50.0%）	4（40.0%）	
No	20（100.0%）	21（65.6%）	19（50.0%）	6（60.0%）	
Days of using sanitary pads					0.007
≥10 days per month	2（10.0%）	10（31.3%）	9（23.7%）	7（70.0%）	
＜10 days per month	18（90.0%）	22（68.7%）	29（76.3%）	3（30.0%）	
Smoking or not					0.710
Yes	0	2（6.3%）	1（2.6%）	0	
No	20（100.0%）	30（93.7%）	37（97.4%）	10（100.0%）	
Age of first sex					0.152
＜20 years old	1（5.0%）	9（28.1%）	5（13.2%）	1（10.0%）	
≥20 years old	19（95.0%）	23（71.9%）	32（86.8%）	9（90.0%）	
Number of sex partner					0.047
1	18（90.0%）	15（46.8%）	21（55.3%）	7（70.0%）	
2	2（10.0%）	11（34.4%）	13（34.2%）	3（30.0%）	
≥3	0	6（18.8%）	4（10.5%）	0	
Frequency of sex					0.025
＜ 1 time per week	16（80.0%）	20（62.5%）	20（52.7%）	2（20.0%）	
2-3 times per week	4（20.0%）	9（28.1%）	14（36.8%）	8（80.0%）	
≥ 3 times per week	0	3（9.4%）	4（10.5%）	0	
Methods of contraception[§]					0.001
None	2（10.0%）	2（6.3%）	5（13.2%）	2（20.0%）	
Condom	15（75.0%）	12（37.5%）	13（34.2%）	1（10.0%）	
Oral contraceptive	0	1（3.1%）	2（5.3%）	0	
Intrauterine device	3（15.0%）	3（9.4%）	6（15.8%）	5（50.0%）	
Others	0	14（43.7%）	12（31.6%）	2（20.0%）	
Parity					0.955

≤1 time	18（90.0%）	30（93.8%）	35（92.1%）	9（90.0%）	
≥2 times	2（10.0%）	2（6.2%）	3（7.9%）	1（10.0%）	
Number of abortion					0.222
≤1 time	12（60.0%）	27（84.4%）	30（78.9%）	7（70.0%）	
≥2 times	8（40.0%）	5（15.6%）	8（21.1%）	3（30.0%）	
History of vaginitis					0.002
Yes	2（10.0%）	14（43.7%）	21（55.3%）	7（70.0%）	
No	18（90.0%）	18（56.3%）	17（44.7）	3（30.0%）	

Abbreviations: hrHPV: high-risk human papilloma virus; TCT: thin-prep cytology test; Ascus: atypical squamous cell of undetermined significance; LSIL: low-grade squamous intraepithelial lesion; HSIL: high-grade squamous intraepithelial lesion; ASC-H: atypical squamous cell- cannot exclude HSIL.

*: *P* value were calculated by ANOVA , Chi-square test or Fisher exact probability test.

†: “-” means no sense; ‡: HPV subtypes might overlapped in each participant; §: Methods of contraception means the most common way of birth control.

Figures

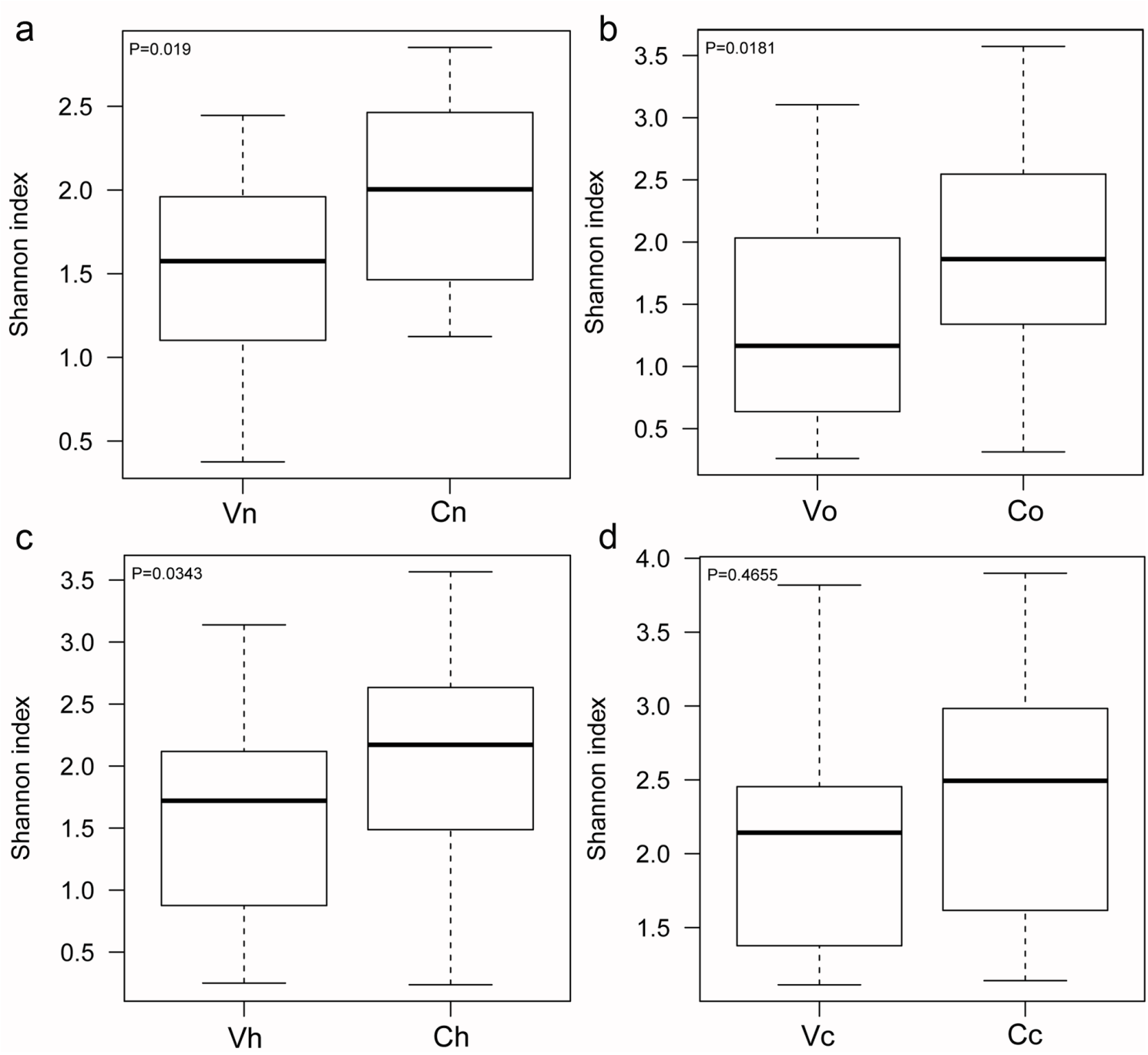

Figure 1

Alpha diversity of vaginal and cervical microbiota. The greater the Shannon value, the higher the diversity of community. (a, group N, n=20, P=0.019. b, group O, n=32, P=0.0181. c, group H, n=38, P=0.0343. d, group C, n=10, P=0.4655)

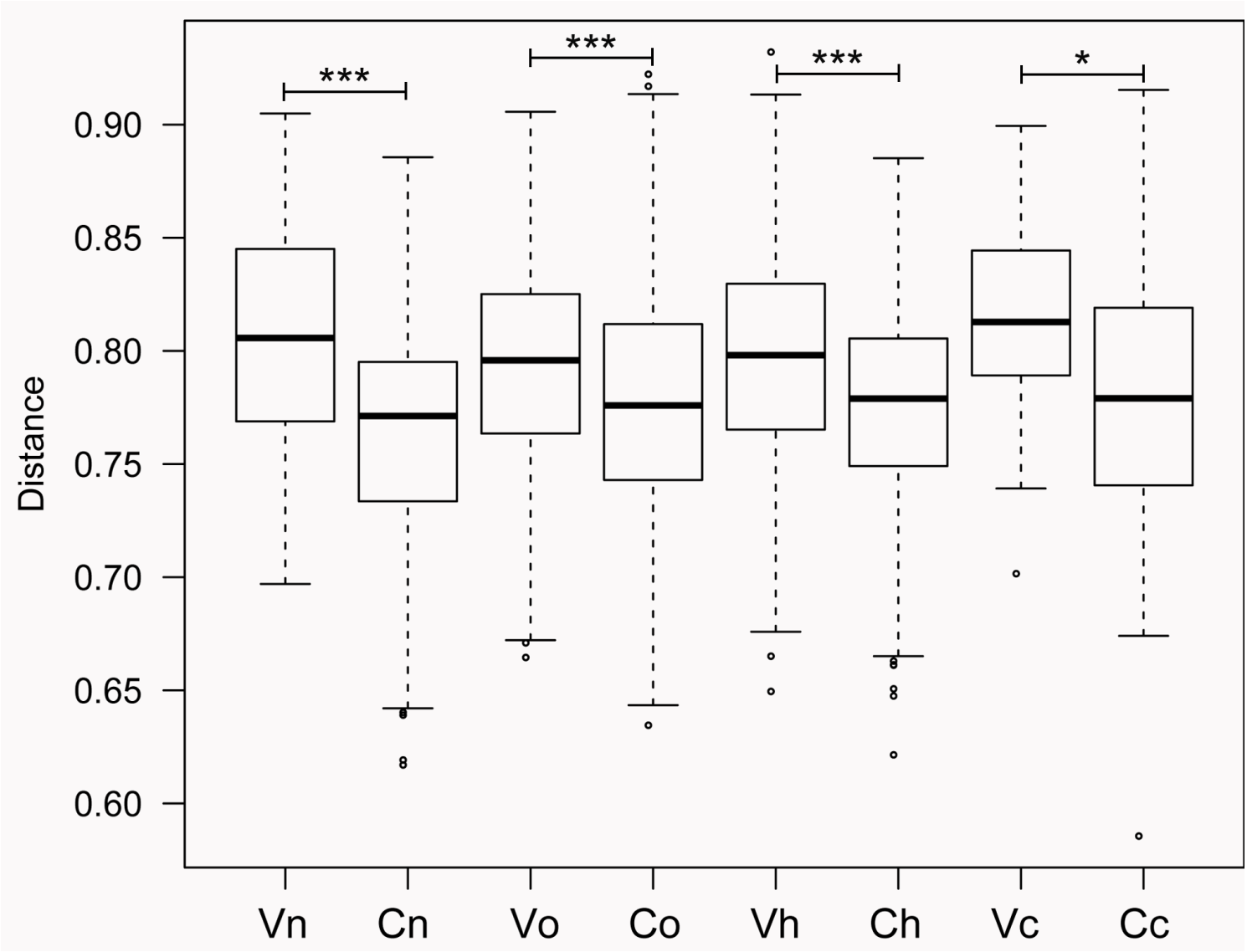

Figure 2

Beta diversity of vaginal and cervical microbiota. Unifrac distance reflects evolutionary distances of strains, ranging from 0 to 1. (*, P ≤0.05; **, P ≤0.01; ***, P ≤0.001)

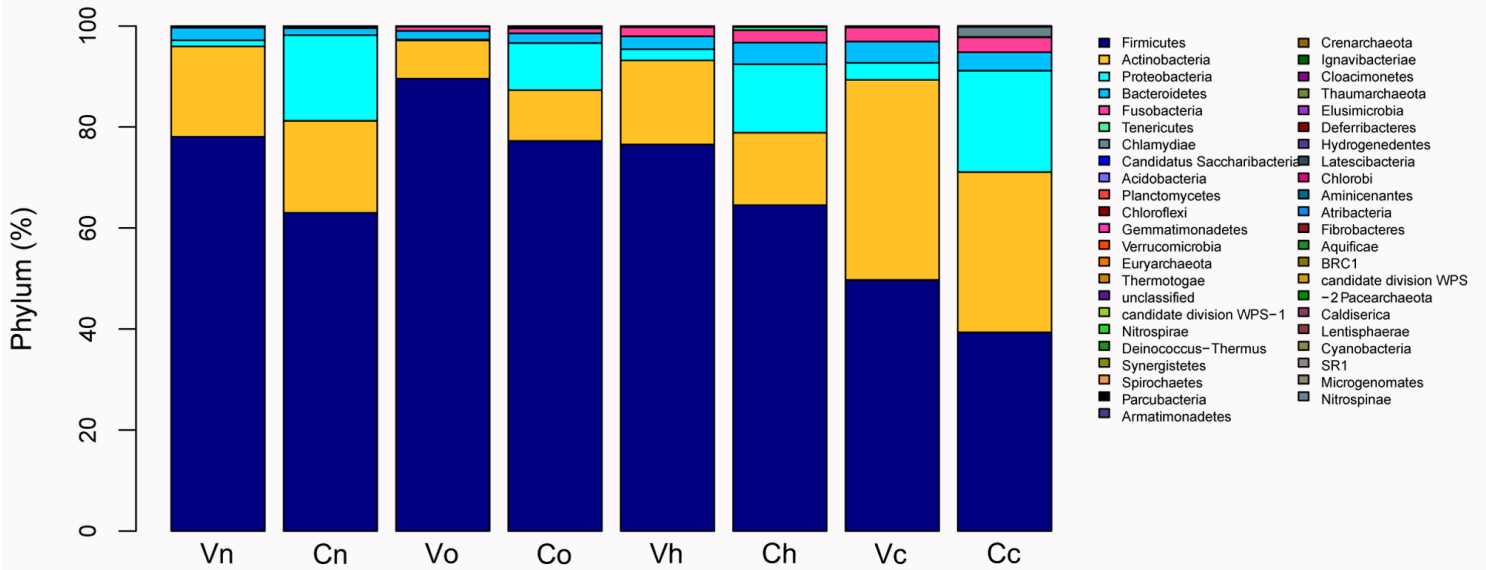

Figure 3

Vaginal and cervical microbiota distribution at phylum level. The abscissa denotes groups and ordinate denotes percentage of microbes at phylum level.

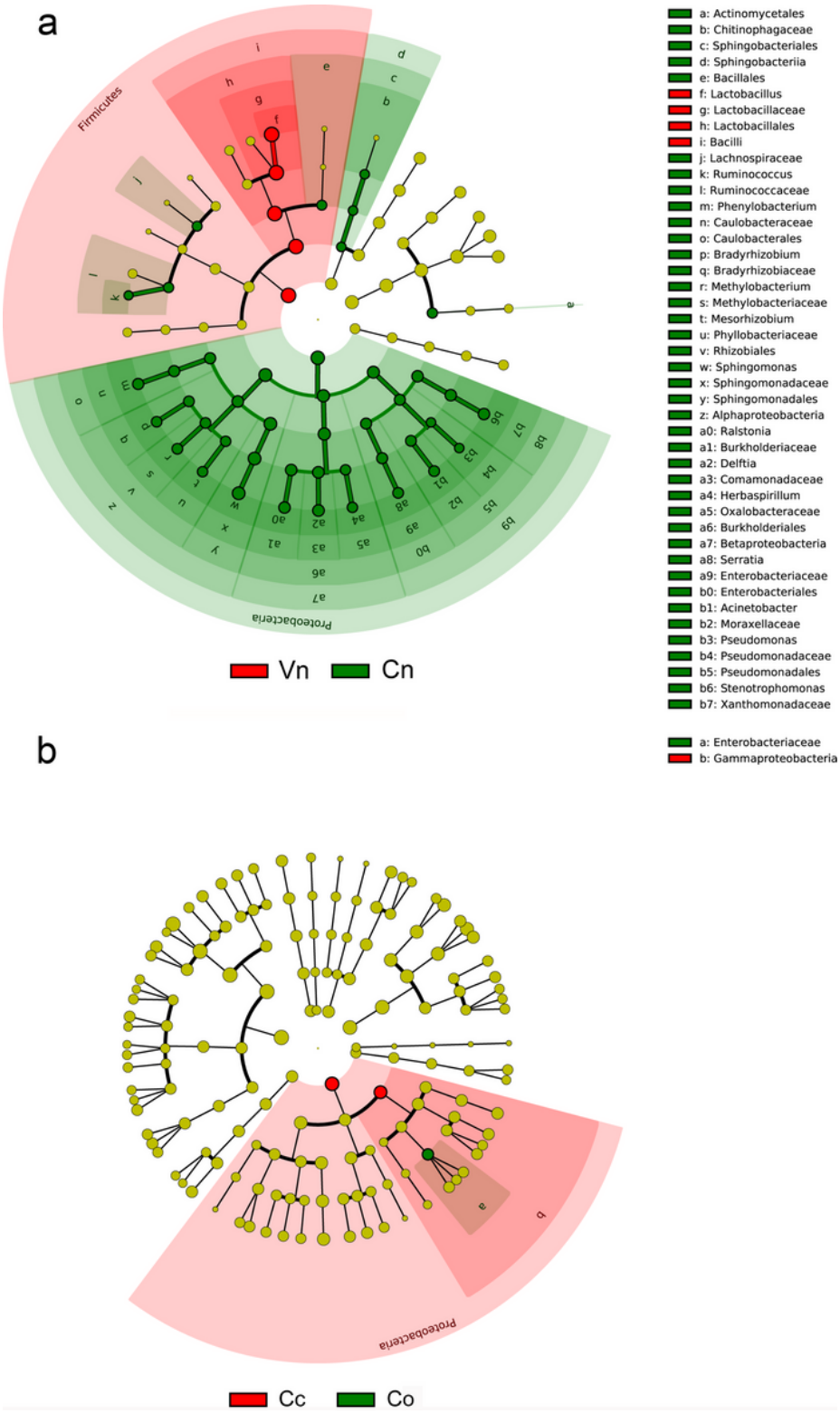

Figure 4

LEfSe linear discriminant analysis of vaginal and cervical microbiota. In LEfSe cladogram, different color in the branches represent microbes that associated with relevant group. (a, Vn vs Cn. b, Comparison of Vn, Vo, Vh, Vc, Cn, Co, Ch and Cc)

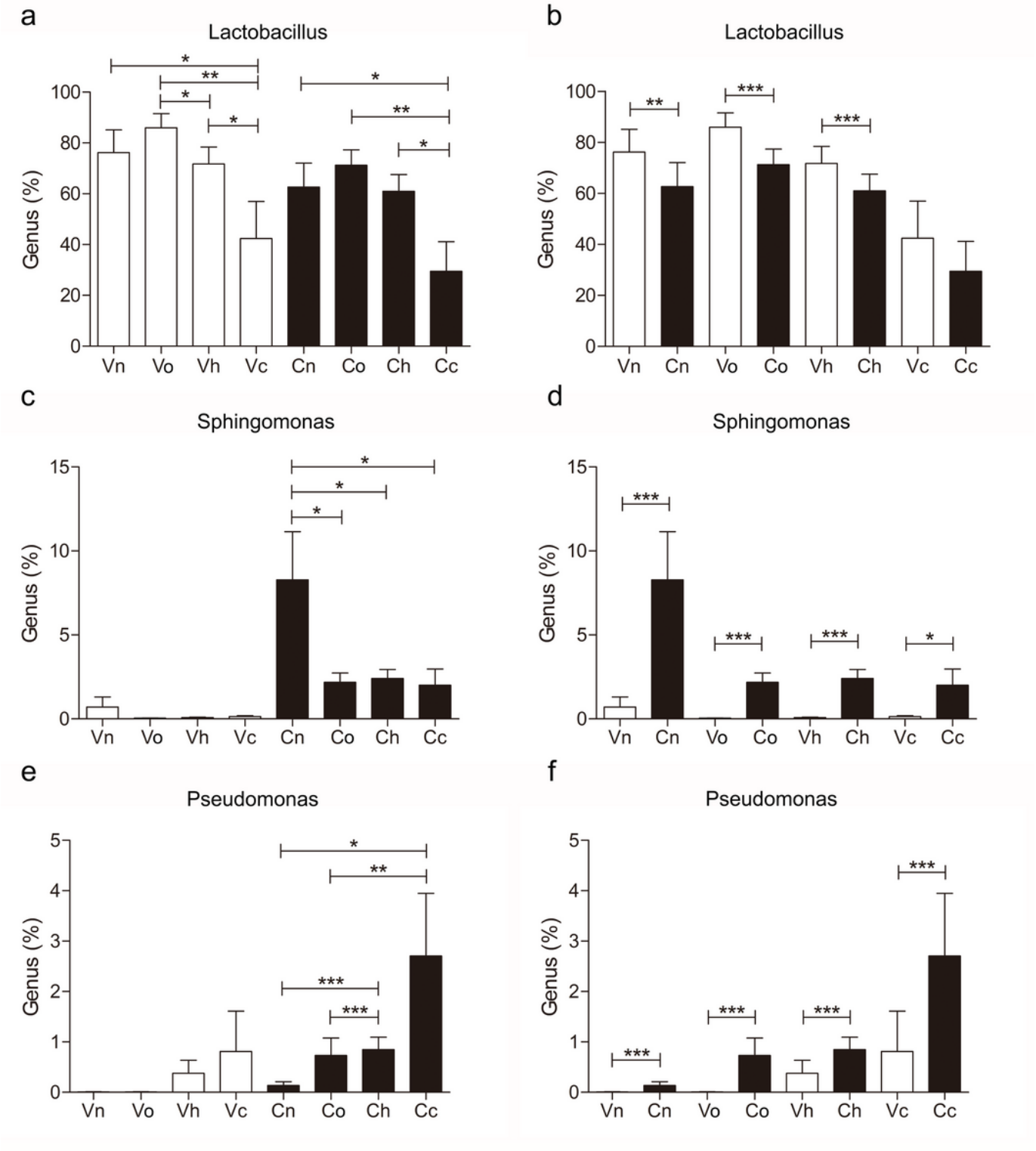

Figure 5

Distribution of Lactobacillus, Sphingomonas and Pseudomonas in the vagina and cervix of the subjects. (*, P ≤0.05; **, P ≤0.01; ***, P ≤0.001)

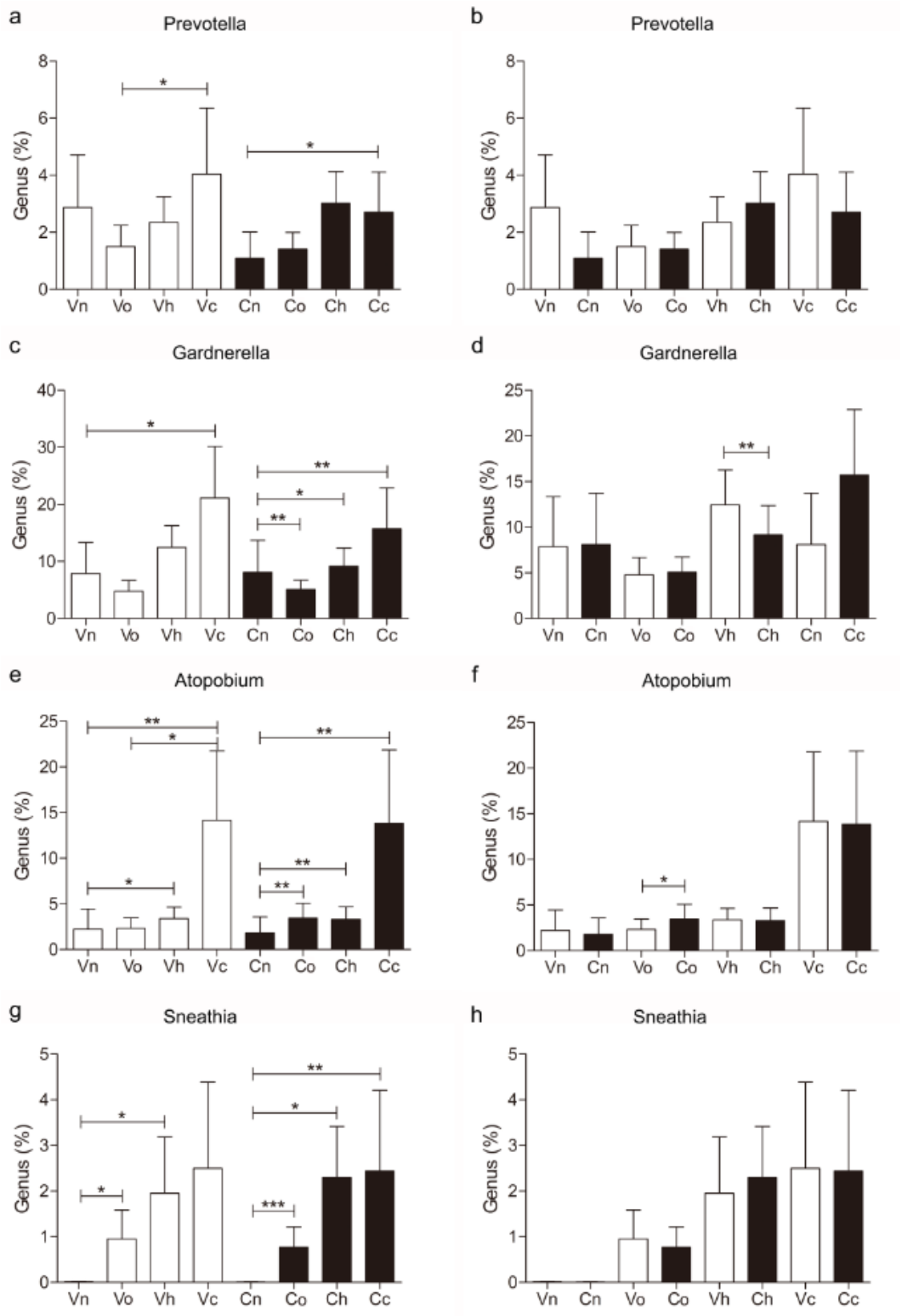

Figure 6

Distribution of Prevotella, Gardnerella, Atopobium and Sneathia in the vagina and cervix of the subjects. (*, P 0.05; **, P 0.01; ***, P 0.001)

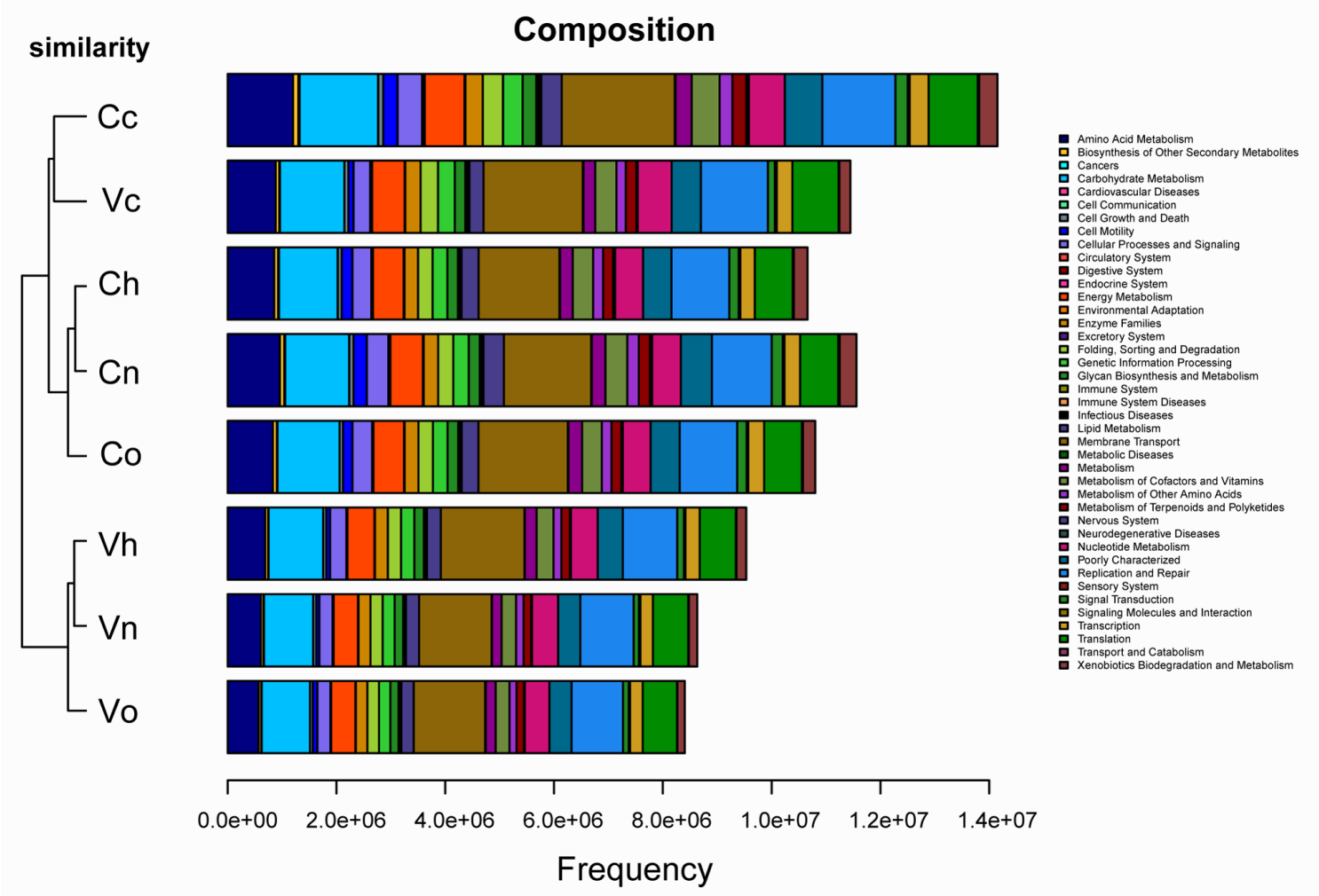

Figure 7

Functional prediction of vaginal and cervical microbiota. The cluster tree and functional bar plot are combined. The sample cluster tree based on bray-curtis is shown on the left in which the length of the branch represents the evolutionary distance between groups. Horizontal bars represent functional composition and abundance of microbiota in each group.