Original Article

Microbiome-metabolome analysis reveals cervical lesion alterations

Hanjie Xu[1,†], Lou Liu[1,†], Feng Xu[1,†], Min Liu[1,†], Yuexiao Song[1], Jiale Chen[1], Huiying Zhan[1], Ye Zhang[1], Dexiang Xu[2], Yu Chen[3,*], Mudan Lu[1,*], and Daozhen Chen[1,4,*]

[1]Department of gynaecology, the Affiliated Wuxi Maternity and Child Health Care Hospital of Nanjing Medical University, Wuxi 214000, China, [2]Department of Toxicology, Anhui Medical University, Hefei 230032, China, [3]Research Institute for Reproductive Health and Genetic Diseases, the Affiliated Wuxi Maternity and Child Health Care Hospital of Nanjing Medical University, Wuxi 214000, China, and [4]Department of Gynaecology, Haidong No. 2 People's Hospital, Haidong 810699, China

[†]These authors contributed equally to this work.

*Correspondence address. Tel: +86-13616178275; E-mail: cy-78@hotmail.com (Y.C.) / Tel: +86-18915283123; E-mail: lumudan0527@163.com (M.L.) / Tel: +86-13584189188; E-mail: chendaozhen@163.com (D.C.)

Abstract

Cervical cancer (CC) continues to be one of the most common cancers among females worldwide. It takes a few years or even decades for CC to arise in a minority of women with cervical precancers. An increasing corpus of studies today indicates that local microecology and carcinogenesis are intimately related. To investigate the changes in cericovaginal microecology with the development of cervical cancer, we performed 16S rDNA sequencing and metabolomic analysis in cericovaginal fluid from 10 LSIL patients, 10 HSIL patients, 10 CC patients and 10 healthy controls to reveal the differential flora and metabolites during cervical carcinogenesis. Carcinogenesis is associated with alterations in microbiome diversity, individual taxa, and functions with notable changes in *Lactobacillus*, *Prevotella* and *Aquabacterium*, as well as in cervicovaginal metabolites that correlate with cervicovaginal microbial patterns. Increased bacterial diversity and a decline in the relative abundance of *Lactobacillus*, the dominant species in the cericovaginal flora, are observed when cervical lesions advance. According to KEGG pathway enrichment analysis, lipids and organic acids change as cervical cancer progresses, and the phenylalanine, tyrosine, and tryptophan biosynthesis pathway is essential for the development of cervical cancer. Our results reveal that microbic and metabolomic profiling is capable of distinguishing CC from precancer and highlights potential biomarkers for the early detection of cervical dysplasia. These differential microorganisms and metabolites are expected to become a potential tool to assist in the diagnosis of cervical cancer.

Key words cervical cancer, cervical squamous intraepithelial lesion, metabolomics, microbiome

Introduction

Cervical cancer (CC) continues to be the second leading cause of cancer death among females aged 20 to 39 years [1]. According to the Global Cancer Observatory 2018 database, the estimated morbidity of cervical cancer is 0.0131% globally and 0.0107% in China [2]. In contrast to the decreasing trend of incidence in developed countries, the CC incidence has gradually risen in China [3]. Over 44 million females will be diagnosed with CC in the following 50 years if the primary and secondary prevention programmers are not well implemented in low- and middle-income areas [4].

Epidemiological studies have established that the persistence of high-risk types of human papillomavirus (HPV) infection is the predominant cause of CC and its precursor lesions [5–7]. Most HPV infections can be cleared spontaneously by the immune system. Only a few persist for years and ultimately lead to cancer [8]. Cervical lesions, including low-grade squamous intraepithelial lesion (LSIL) and high-grade squamous intraepithelial lesion (HSIL) can develop into CC within approximately 10 to 20 years [9]. The HPV test and Pap smears are utilized to determine the risk of CC as a

routine screening tool. However, they are not convenient enough. If the LSIL state exists for a long time or cytology abnormality occurs after loop electrosurgical excision procedure (LEEP), the confidence of patients would be affected. The emergence of the cervicovaginal microbiome and metabolome may help us improve the prediction of CC.

Researchers have found that HPV infection is associated with vaginal microecological imbalances, including bacterial vaginosis, trichomoniasis, vulvovaginal pseudomonal yeast disease, aerobic vaginitis and mixed infections [10–12]. It is now generally accepted that vaginal microecology is significantly associated with HPV infection, cervical lesions and cervical cancer, and plays a significant role in the course of cervical lesions [13], but the exact mechanism is not yet clear. Vaginal microecological imbalance induces cervical cancer probably through the metabolism of pathogenic flora that affects the host's immune barrier, the growth of pathogenic flora that causes damage to the cervical mucosa and activates the expressions of inflammatory factors, and the *disease-causing bacteria* may lead to hypoxia and low pH and affect the clearance of HPV and may facilitate the cervicitis-cervical cancer transition [14–16]. The healthy vaginal microecology is dominated by *Lactobacillus*, which plays an important role in maintaining the vaginal microecological balance through acid production, H_2O_2 production, competitive adhesion, secretion of antibacterial peptides, and induction of local immune responses [17–19]. Under balanced vaginal microecological conditions, *Lactobacillus* may exert anti-HPV and anti-inflammatory effects, whereas under imbalanced vaginal microecological conditions, it may induce further progression of CIN to cervical cancer. The correlation between vaginal microecology and cervical lesions may play an important role in the early warning and intervention of cervical cancer. It is generally accepted that bacterial diversity and composition are altered in HPV infection and cervical lesions, however the changes in cervicovaginal flora and their metabolism remain unclear during carcinogenesis among Chinese women.

In the current study, we performed 16S rDNA sequencing and metabolomics analysis to identify potential biomarkers in cervical carcinogenesis.

Materials and Methods
Study subjects and sample collection
Participants were consecutively recruited at the Affiliated Wuxi Maternity and Child Health Care Hospital of Nanjing Medical University. For exploratory studies, it is not necessary to calculate the sample size. To estimate the sample size needed, we calculated the sample size by one-way analysis of variance F tests using effect size via PASS (Power Analysis and Sample Size) software. To achieve a power > 0.8 and P value < 0.05, the sample size was 19 per group when f = 0.4. Because the sample size for a pilot study is approximately 10% ~ 20% of the total sample size of the study, we chose 10 per group in our discovery study. For this, we believe the sample size in our study is adequate as a discovery study. Forty nonpregnant women diagnosed with cervical dysplasia and CC, as well as healthy HPV-negative women, were enrolled and contributed to the study. Colposcopy-directed biopsy (CDB) and cytology were used to classify patients into groups. All cases were diagnosed by two independent doctors according to clinical and pathological features. Overall, patients were divided into 4 groups (n = 10 each): women with LSIL, women with HSIL, women with

CC, and healthy participants as controls. The details of the exclusion criteria are listed in Supplementary Table S1. A single-use vaginal speculum was inserted with lubricant, and two cervicovaginal swabs were collected by a clinician. Following collection, the swabs were immediately placed in liquid nitrogen for flash freezing and then stored at –80°C until analysis.

Ethics approval
This study was approved by the Institutional Review Boards of The Affiliated Wuxi Maternity and Child Health Care Hospital of Nanjing Medical University (approval number: 2020-01-0309-06), and written informed consent was obtained from each participant. The study has been registered with the Chinese Clinical Trials Registry (registration number: ChiCTR2000034596, reg date: 2020/7/11). All methods were performed in strict accordance with the relevant guidelines and regulations.

HPV DNA testing
Hybrid Capture 2 (HC2) technology was applied for HPV DNA testing, which can differentiate between 2 HPV DNA groups, most-high-risk (HPV types 16, 18) and high-risk types (HPV types 31, 33, 35, 39, 42, 43, 44, 45, 51, 52, 56, 58, 59, and 68). The participants enrolled had all been infected with at least one HPV type above, and all were HPV negative at the time of resampling. The effect of HPV on vaginal microecology was therefore excluded.

Cervicovaginal DNA extraction for microbiome analysis
The cervicovaginal swab was placed in TENS buffer (5 M sodium chloride, 10% Triton X-100, 1 M Tris-HCl, pH 8.0, EDTA) containing 10% SDS and 20 mg/mL proteinase K, and then incubated overnight at 55°C. Proteins were removed by phenol/chloroform/isoamyl alcohol extractions, and the DNA was precipitated with isopropanol the following day. After wash with 75% ethanol, DNA was resuspended in TE buffer. DNA was quantified with a Qubit Fluorometer by using a Qubit dsDNA BR Assay kit (Invitrogen, California, USA), and the quality was checked by 1% agarose gel electrophoresis.

Library construction
Library construction was performed on an Illumina MiSeq platform (BGI, Shenzhen, China) as described previously [20]. Briefly, variable regions V3–V4 of the bacterial 16S rDNA gene were amplified with degenerate PCR primers 341F (5′-ACTCCTACGGGAGGCAGCAG-3′) and 806R (5′-GGACTACHVGGGTWTCTAAT-3′).

Sequencing and bioinformatics analysis
Raw reads were filtered to remove adaptors and low-quality and ambiguous bases, and then paired-end reads were added to tags by the Fast Length Adjustment of Short reads program (FLASH, v1.2.11)[21] to obtain the tags. The tags were clustered into ASVs using DADA2 [22]. Then, ASV representative sequences were taxonomically classified using Ribosomal Database Project (RDP) Classifier v.2.2 with a minimum confidence threshold of 0.6 and trained on the Greengenes database QIIME2 [23].

Statistical analysis of 16S rDNA sequencing data
Alpha and beta diversity were estimated by MicrobiomeAnalyst [24] at the ASV level. Sample clustering was conducted by QIIME2 based on UPGMA. Species accumulation curves and partial least-squares

discrimination analysis (PLS-DA) were plotted with MicrobiomeAnalyst. A GraPhlan map of species composition was created using GraPhlAn. Significant species or functions were determined using R (v3.4.1) based on the Wilcoxon test or Kruskal-Wallis test.

Metabolite extraction

The samples were thawed slowly at 4°C, and metabolites were extracted using 300 μL of methanol:acetonitrile (2:1, v:v) at –20°C for 2 h. The samples were then centrifuged at 4000 g for 20 min at 4°C. After centrifugation, 300 μL of supernatant was dried with a vacuum concentration meter and then redissolved in 150 μL methanol:H$_2$O (1:1, v:v). Then, the samples were centrifuged at 4000 g for 30 min at 4°C. Ten microliters of supernatant was mixed as a quality control (QC) sample.

LC-MS analysis of metabolites

Cervicovaginal samples were eluted from swabs using methanol and subjected to nontargeted liquid chromatography-mass spectrometry (LC-MS)-based metabolomics conducted by BGI Tech Solutions Co., Ltd (Shenzhen, China). Briefly, LC-MS was performed using ultra-performance liquid chromatography (Waters 2D UPLC; Waters, Milford, USA) and high-resolution mass spectrometry (Thermo Fisher Scientific, Waltham, USA), and samples were run in positive and negative ion sense. To provide more reliable experimental results, random sorting of samples was carried out to reduce systematic errors. The QC sample was inserted for every 10 samples.

Chromatographic separation was performed using a BEH C18 column (1.7 μm particle size, 2.1 mm × 100 mm inner diameter; Waters). For the positive ion mode, mobile phase A was a 0.1% formic acid aqueous solution, and mobile phase B was 100% methanol containing 0.1% formic acid. For the negative ion mode, mobile phase A was an aqueous solution containing 10 mM ammonia formate, and mobile phase B was 95% methanol containing 10 mM ammonia formate. The following gradient was used for elution: 0–1 min, 2% mobile phase B; 1–9 min, 2%–98% mobile phase B; 9–12 min, 98% mobile phase B; 12–12.1 min, 98%–2% mobile phase B; 12.1–15 min, 2% mobile phase B. The flow velocity was 0.35 mL/min, the column temperature was 45°C, and the injection volume was 5 μL.

A Q Exactive HF mass spectrometer (Thermo Fisher Scientific) was used to collect primary and secondary mass spectrometry data. The scanning range of mass-to-charge ratio (m/z) was 70–1050, the primary resolution was 120,000, the AGC was 3e6, and the maximum injection time was 100 ms. According to the strength of the parent ion, Top3 was selected for fragmentation, and secondary information was collected. Secondary resolution was 30,000, AGC was 1e5, maximum injection time was 50 ms, and stepped nce was 20, 40 and 60 eV. The parameter settings of the ion source (ESI) were as follows: sheath gas flow rate, 40; aux gas flow rate, 10 mL/min; spray voltage (|KV|), 3.80 in ESI$^+$ and 3.20 in ESI$^-$; capillary temperature, 320°C; and aux gas heater temperature, 350°C.

Data processing and metabolite identification

Features in the LC–MS metabolomics raw data were aligned, and peak areas were determined using XCMSonline (https://xcmsonline.scripps.edu/) [25]. The method of feature detection was centWave. The peak width range was set from 5 to 20. The method of retention time correction was determined. Alignment parameter:

mzwid was 0.025, bw was 5. Metlin (https://metlin.scripps.edu/) [26] was used to identify the individual m/z features. Search for isotopes$^+$ adducts for annotation. Annotation parameter: ppm was 5 m/z absolute error was 0.015. Identification parameter: ppm was 15, adducts were [M-H]$^-$, [M-H$_2$O-H]$^-$, [M-Na-2H]$^-$, [M-Cl]$^-$, [M-K-2H]$^-$, [M-FA-H]$^-$ and [M-2H]$^{2-}$ for ESI$^-$ and [M + H]$^+$, [M + NH$_4$]$^+$, [M + Na]$^+$, [M + H-H$_2$O]$^+$, [M + K]$^+$, [M + 2Na-H]$^+$, [M + 2H]$^{2+}$ for ESI$^+$. The identified metabolites were annotated with the HMDB, KEGG, and NCBI pubchem databases. Raw data files can be accessed through the Metabolights open access database (Study ID MTBLS1971).

Statistical analysis of metabolomics data

Statistical analysis was performed by an online tool called MetaboAnalyst 4.0 (http://www.metaboanalyst.ca) [27]. Hierarchical clustering was used to detect the classification ability and concentration levels of metabolites. Log transmission and autoscaling were used on metabolomics data to make the data obey a normal distribution for drawing a heatmap. The Kruskal–Wallis test of the original data was used for pairwise comparisons. Chemometrics analysis of principal component analysis (PCA) and partial least squares-discriminant analysis (PLS-DA) were applied to reveal the global metabolic difference of different groups and evaluate the stability and credibility. The cross validation (CV) method was used to estimate the predictive ability of the model. Pathway information was extracted from KEGG (www.kegg.jp/kegg/kegg1.html) [28]. GRAPHPAD PRISM version 8.3 (GraphPad, Inc., San Diego, USA) was used for statistical analysis and generation of figures, with statistical significance achieved when P value is < 0.05.

Results

Participant characteristics

A total of 40 cervicovaginal samples were collected from patients in this study. The average age of the participants was 48.28 years (SD: 12.76; range: 18–78). The mean ages calculated for the healthy control (HC), LSIL, HSIL and CC groups were 35.6, 54.7, 44.8 and 58.0 years, respectively. Patients with high-risk HPV infection were selected from our sample, including HPV16, 18, 31, 33, 56, 58, 59, 68, 35, 39, 45, 51 and 52. The clinical characteristics of the participants are presented in Supplementary Table S2.

Increased bacterial diversity in cervicovaginal microbiota associated with cervical carcinogenesis

To explore the changes in vaginal flora during the development of cervical lesions, 16S rDNA sequencing was performed. A total of 1,790,607 high-quality 16S rDNA reads were obtained in our microbiome investigation, with a mean read count of 47,121.24 (range 18,752–64,758) per sample. After taxonomic assignment, 2059 ASVs were obtained. The rarefaction curve of all samples showed that the sampling efforts were adequate (Supplementary Figure S1 and Table S3). At the phylum and genus levels, the relative proportions of dominant taxa were assessed by microbial taxon classification in each group (Figure 1A). The stacking bar charts indicate variability in cervicovaginal microbiota across samples in each group, especially in the CC group.

To evaluate the differences in bacterial diversity between the four groups, sequence alignment was performed to estimate alpha diversity and beta diversity (Figure 1B,C). The differences in the Shannon (P = 0.04) and Simpson (P = 0.02) indexes were statisti-
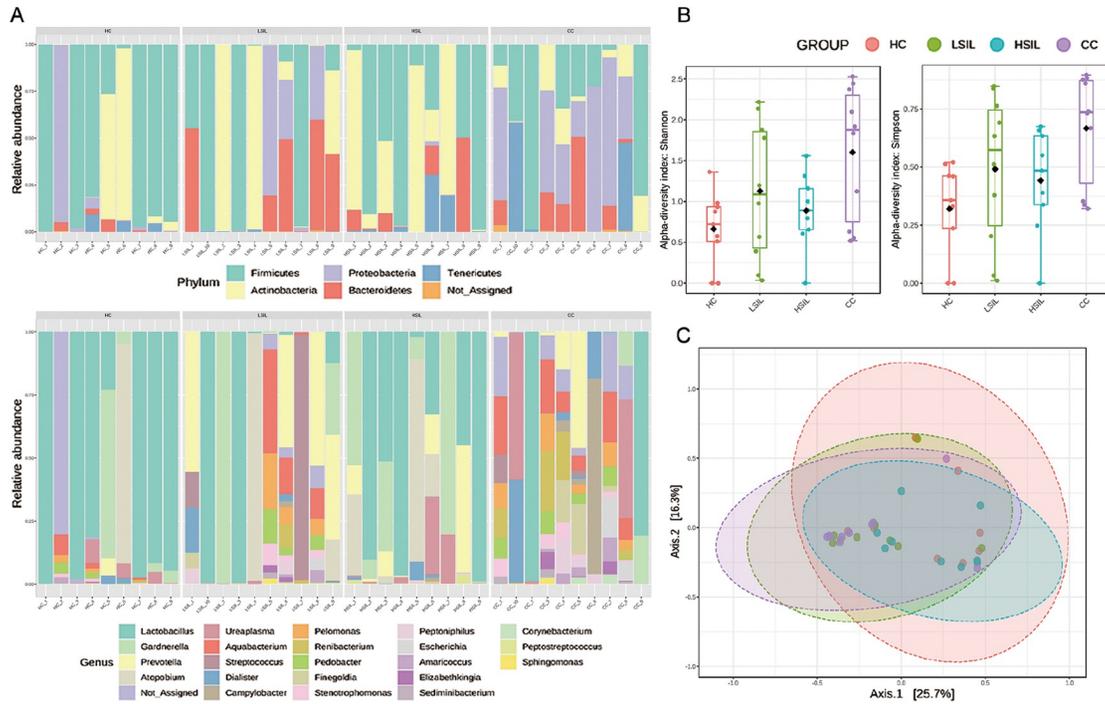
Figure 1. Cervicovaginal microbiome structure and diversity analysis (A) Stacked Bar Charts of different groups representing the relative abundance of bacterial phylum (upper) and genus (lower). (B) Species α diversity differences between the 4 groups were estimated by Shannon (left) and Simpson (right) indices. *$P < 0.05$. (C) Principal coordinates analysis (PCoA) plots using Bray-Curtis index.

cally significant between the HC and CC groups (Figure 1B). In Figure 1C, the PCoA plots revealed a separation trend between the 4 groups ($F = 1.8557$, $R^2 = 0.1407$, $P = 0.008$). These results suggested that the bacterial diversity was increased and the composition was changed by the influence of cancer carcinogenesis.

Cervical cancerization is associated with specific cervicovaginal microbial taxa

To identify individual microbial taxa associated with cervical carcinogenesis, we assessed the abundances of each feature among the four groups. The microbial association mapping results are illustrated on a taxonomic tree (Figure 2) using GraphlAn [29]. Cervical cancerization is associated with the microbial genera *Gardnerella* (B), *Prevotella* (D), and *Lactobacillus* (F), which are clustered in the families *Bifidobacteriaceae* (A), *Prevotellaceae* (C) and *Lactobacillaceae* (E), respectively. At the *phylum* level, *Actinobacteria* exhibited a decline only in cervical cancer patients, *Proteobacteria* declined only in HSIL patients, while *Bacteroidetes* exhibited a rise only in LSIL patients (Figure 2). Compared to that in HC, the relative abundance of *Lactobacillus* was significantly reduced in LSIL ($P = 0.036$) and CC ($P = 0.023$) at the *genus* level. This suggests that there may be a negative correlation between the abundance of *Lactobacillus* and the progression of cervical cancer. In contrast, the relative abundance of *Prevotella* was significantly elevated in LSIL ($P = 0.015$) and had a tendency for a gradual decline with the progression of cervical lesions (Figure 2). *Aquabacterium* appeared to be a specific species that was increased only in LSIL and CC. Additionally, *Lactobacillus* showed a negative correlation with other *genera* (Figure 2). From these results, it can be inferred that there are changes in the vaginal flora during cervical cancer progression and that there are specific species corresponding

to each stage of change.

Overall metabolomics analysis of cervicovaginal fluid samples

LC-MS-based metabolomics was applied to detect small molecule metabolites in vaginal discharge to reveal an overview of metabolism in vaginal microecology. Representative LC-MS base peak chromatograms (BPCs) of the identified compounds are presented in Figure 3. The total ion chromatograms for all the participants are shown in Figure 4A,B. Although the overall signals of the four groups were similar, each group seemed to have definite signals. After peak alignment and missing value removal, 16927 ESI$^+$ and 8672 ESI$^-$ features were obtained. Statistical analyses by PCA showed separation between the four groups (Figure 4C,D). The PCA plots, including QC samples, guaranteed the reliability of the experimental results. The wide distribution of the samples in PCA was due to different degrees of metabolic manifestations *in vivo* and/or by the subsequent sampling date. Another reason may be that the development of cervical cancer is a gradual process, and the metabolites also change gradually.

Altered metabolites in cervical carcinogenesis

To explore the changes in metabolites during cervical lesions, bioinformatics analysis was further carried out. There were 155 differential metabolites ($P < 0.05$) identified among the 4 groups (Figure 5A). Based on the 155 differential metabolites, PLS-DA was applied to sort out components that are responsible for differentiation among various statuses ($R^2 = 0.95686$, $Q^2 = 0.70423$). As shown in Figure 5B, groups HC and CC were segregated clearly from each other, indicating that these selected metabolites can better distinguish different groups. To determine unique metabolites for
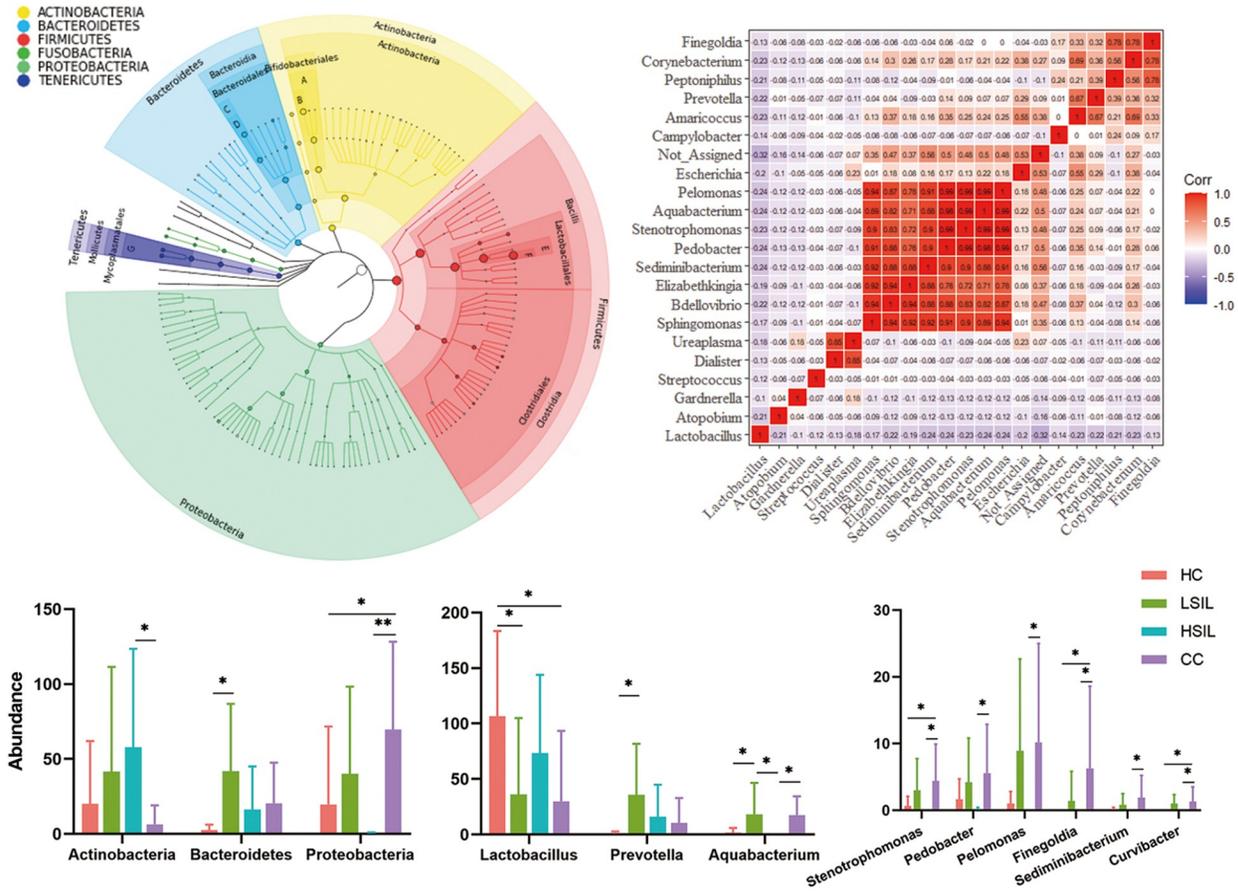
Figure 2. Specific taxa associated with cervical cancerization Taxonomic tree generated using Graphlan (upper left). The nodes on the tree from inner to outer circles are the phylum, class, order, family, genus, and species rank. The highlighted nodes *Bifidobacteriaceae* (A), *Gardnerella* (B), *Prevotellaceae* (C), *Prevotella* (D), *Lactobacillaceae* (E) and *Lactobacillus* (F) are vital species among cervical cancerization. Correlation analysis by Spearman's rank correlation coefficient (upper right). Barplot of specific taxa under phylum and genus level (lower). $*P < 0.05$, $**P < 0.01$.

different phenotypic groups, pairwise comparisons were performed (Figure 5C). Lipids were significantly decreased in LSIL ($P = 0.042$) and HSIL ($P = 0.042$) compared to those in HC, and significantly elevated in CC compared to those in LSIL ($P = 0.004$) and HSIL ($P = 0.005$). To explore the metabolic pathways that potentially contribute to cervical cancer progression, we performed KEGG pathway enrichment analysis using the list of 155 differential metabolites and found that the phenylalanine, tyrosine and tryptophan biosynthesis pathways were crucial in cervical carcinogenesis (Figure 5D). These data indicated that there are corresponding changes in the body's metabolism as the cervical lesion progresses, in addition to changes in the vaginal flora.

## Associations of cervicovaginal microbial species with cervicovaginal metabolites

Correlation analysis was performed to discover the association between vaginal flora and metabolites altered in cervical lesions. The Spearman's rank correlation between cervicovaginal genera and metabolites is shown in Figure 6. Control-enriched *Lactobacillus* was positively correlated with lipids such as PC(20:2 (11Z,14Z)/14:0) and was negatively correlated with organic acids such as tyrosine, busulfan and N-palmitoyl phenylalanine. Conversely, *Prevotella* was positively correlated with organic acids such as N-palmitoyl phenylalanine. These data indicated that vaginal

dysbiosis may influence the progression of cervical lesions by regulating body metabolism.

Discussion

Emerging evidence indicates the role of cervicovaginal flora and metabolites in the pathogenesis of cervical cancer. To our knowledge, this is the first report characterizing the cervicovaginal bacteria and metabolites of cervical cancer and precancer in China. This study aimed to explore the relationship between the flora and metabolites of vaginal microecology and the progression of cervical cancer.

Nejman *et al.* [30] showed changes in the abundance and diversity of flora in the microecology of a variety of tumours, such as breast, lung, ovarian and pancreatic cancers. We found that the microbial diversity was significantly increased in CC (Figure 1B), which is consistent with previous studies [31,32]. Currently, the vaginal microbiome is classified into 5 major groups, named community state type (CST). Four CSTs are dominated by lactic acid-producing *Lactobacillus* species, including *Lactobacillus crispatus*, *Lactobacillus gasseri*, *Lactobacillus iners* and *Lactobacillus jensenii*. CST IV is dominated by a low abundance or absence of *Lactobacillus* and a dedicated or parthenogenic anaerobic flora, which is associated with the development of vaginitis [33]. We found that *Lactobacillus* was depleted in the microbiome of cervical
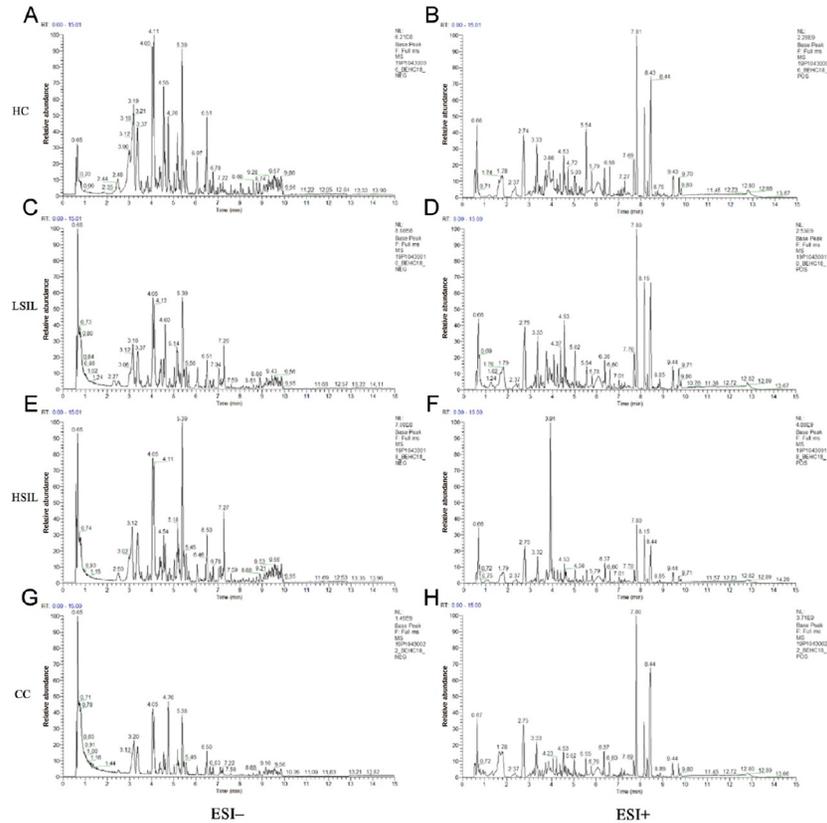
Figure 3. Typical BPC from women with and without cervical lesions The typical base peak chromatogram (BPC) of cervicovaginal fluid in samples from four groups in ESI⁻ (A,C,E,G) and ESI⁺ modes (B,D,F,H). The x-axis represents retention time, and the y-axis represents the charge-to-mass ratio of the features.

cancer patients (Figure 2). In other words, CST IV is associated with cervical cancer progression. We did not explore the association of each CST with cervical cancer due to sample size limitation, but we will explore this part in a future large sample study. Similar to our results, the depletion of *Lactobacillus spp.* was found to be correlated with persistence and slower regression of CIN2 [34,35]. *Lactobacillus spp.* can produce lactic acid to maintain a low pH [36,37] and produce bacteriocin [38,39] to prevent colonization of bacterial vaginosis-associated bacterial species. A *Lactobacillus*-enriched cervicovaginal environment is important for maintaining the cervicovaginal epithelial barrier, preventing HPV infection [40]. When bacterial vaginosis-associated taxa are able to colonize, they produce enzymes and metabolites that are harmful to the barrier and facilitate HPV infection [40]. Conversely, the increase in *Gardnerella* and *Prevotella* appears to be positively associated with cervical cancer progression (Figure 2). Similar to our results, Mykhaylo *et al.* [35] showed that *Gardnerella* is a major biomarker of high-risk HPV progression. *Prevotella spp.* are found in humans and help breakdown proteins and carbohydrates. They can also act as a conditional pathogen, causing periodontitis, enteritis, rheumatoid arthritis, bacterial vaginitis and so on [41–43]. *Prevotella* produces lipopolysaccharides (LPS) and ammonia as part of the vaginal secretions. It is also associated with the production of epithelial cytokines and promotes the growth of other vaginosis-associated bacteria, such as *Gardnerella*, which in turn stimulates *Prevotella*'s growth [43,44]. In conclusion, the progression of cervical cancer is accompanied by an imbalance in the vaginal flora.

We found that metabolism was changed during cervical cancerization, especially lipid and organic acid metabolism (Figure 5). Previous studies revealed that lipid metabolism, such as fatty acid metabolism and sphingolipid metabolism, as well as organic acid metabolism, is altered in the serum of cervical cancer patients [45,46], consistent with our results in cervicovaginal fluid (Figure 5D). Proliferation of cancer cells requires a mass of lipids to make up the membranes and organelles of cancer cells [47]. According to our microbiome-metabolome correlation analysis (Figure 6), *Lactobacillus* was negatively related to tyrosine and positively related to methionine. Consistent with our results, *Lactobacillus* was found to produce tyramine from tyrosine [48–50]. Methionine can increase the number of *Lactobacillus* and attenuate oxidative stress and inflammatory reactions [51]. These studies suggest that changes in microbial composition can alter relevant metabolites. Meanwhile, changes in metabolites influence microorganisms and organic reactions in turn. Therefore, we assume that these flora and metabolites may become potential biomarkers for cervical cancer.

In conclusion, cervical cancer is a serious global health problem. This study used microbiome and metabolomics approaches to explore the molecular changes in cervical cancer-specific flora and metabolites in cervicovaginal secretions. Our findings extend our insights into the relationship between the cervicovaginal microbiota and metabolism during cervical cancerization, pointing to possible future modalities for cancer prevention targeting the cervicovaginal microenvironment.It is a noninvasive approach that may lead to new strategies for the management of women at high risk for cervical
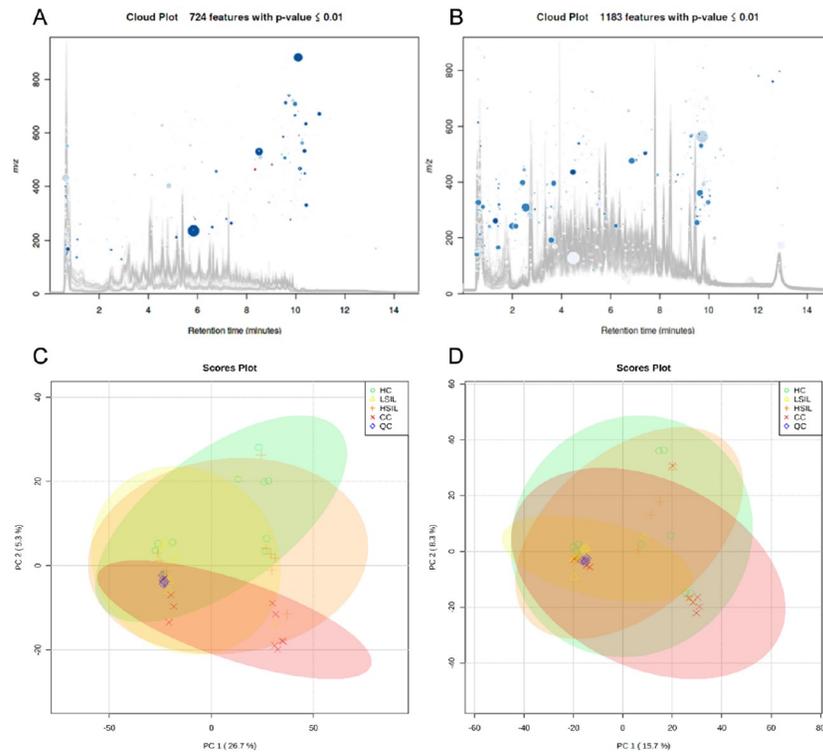
Figure 4. Metabolic profiles of cervicovaginal samples of the ESI+ and ESI– modes   Cloud plots coupled with total ion chromatograms (A&B). After peak alignment and removal of missing values, 8672 electrospray ionization ESI– features (A) and 16927 ESI+ features (B) were obtained. The x-axis represents retention time, and the y-axis represents the charge-to-mass ratio of the features. Each circle in the cloud plot represents 1 differential feature, and the circle size represents the relative concentration of the feature. Metabolic Profiles of cervicovaginal samples including QC samples by PCA (C,D). Score plots of PCA revealed the clustering of samples in negative (left) and positive (right) ion mode. Green, HC; yellow, LSIL; orange, HSIL; red, CC; blue, QC.
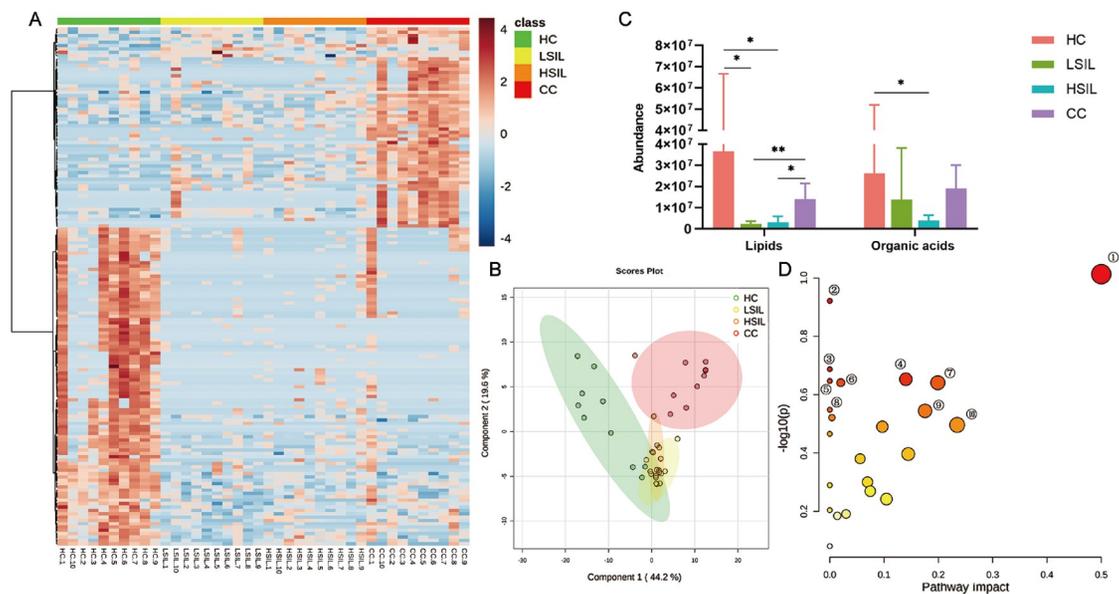
Figure 5. Specific metbabolites associated with cervical cancerization   (A) Heatmap of 155 differential metabolites among cervical carcinogenesis. Study participant identification numbers were provided on the x-axis, and metabolites were listed on the y-axis. (B) PLS-DA 2D scores plot. Each circle represents a sample, and the shadow represents the 95% confidence interval. (C) Barplot of lipids and organic acids among 4 groups. *$P < 0.05$, **$P < 0.01$. (D) KEGG pathway analysis was conducted using Metaboanalyst. X-axis, pathway impart; Y-axis, -log(p). The larger the circle size, the greater the match status. The redder the color, the smaller the P value. (1) Phenylalanine, tyrosine and tryptophan biosynthesis (2) Linoleic acid metabolism (3) Ubiquinone and other terpenoid-quinone biosynthesis (4) Purine metabolism (5) Phenylalanine metabolism (6) Arachidonic acid metabolism (7) Glycerophospholipid metabolism (8) alpha-Linolenic acid metabolism (9) Tyrosine metabolism (10) Glycosylphosphatidylinositol (GPI)-anchor biosynthesis
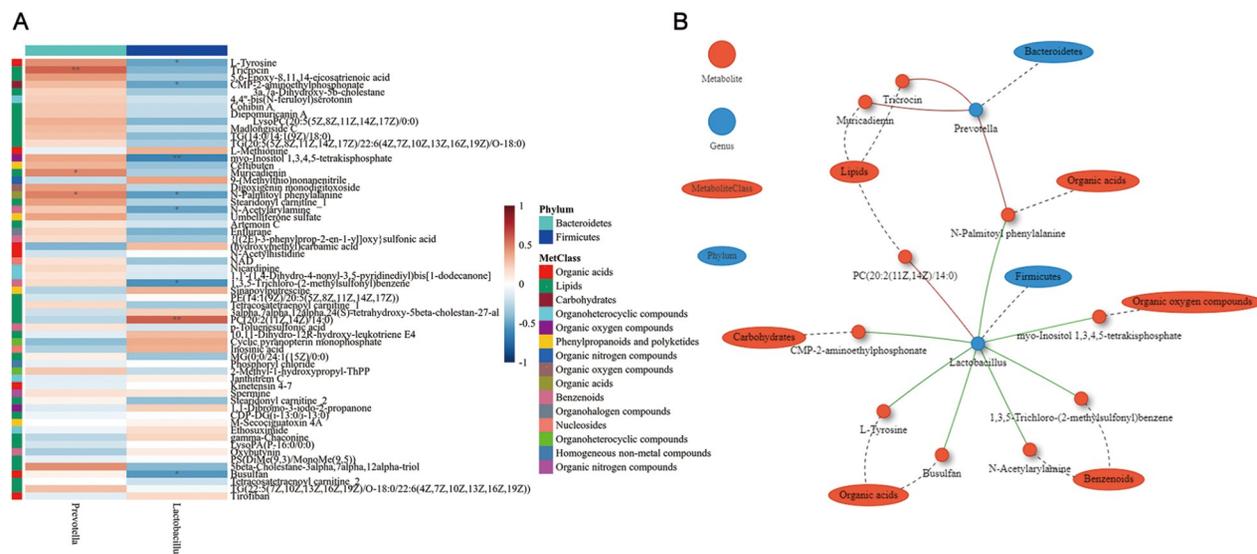
Figure 6. Association of cervicovaginal microbial species with metabolites (A) Heatmap of correlation analysis between cervicovaginal genus and metabolites. Red: positive correlation. Green: negative correlation. (B) Correlation network between cervicovaginal genus and metabolites. Red line: positive correlation. Blue line: negative correlation.

cancer. Nevertheless, the limitation of this study is the relatively small sample size. Large sample data and targeted metabolomics are needed to validate these cervicovaginal biomarkers in the future.

Supplementary Data
Supplementary data is available at *Acta Biochimica et Biphysica Sinica* online.

Conflict of Interest
The authors declare that they have no conflict of interest.

References
1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* 2020, 70: 7–30
2. Arbyn M, Weiderpass E, Bruni L, de Sanjosé S, Saraiya M, Ferlay J, Bray F. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. *Lancet Glob Health* 2020, 8: e191–e203
3. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, *et al.* Cancer statistics in China, 2015. *CA Cancer J Clin* 2016, 66: 115–132
4. Simms KT, Steinberg J, Caruana M, Smith MA, Lew JB, Soerjomataram I, Castle PE, *et al.* Impact of scaled up human papillomavirus vaccination and cervical screening and the potential for global elimination of cervical cancer in 181 countries, 2020–99: a modelling study. *Lancet Oncol* 2019, 20: 394–407

5. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, *et al.* Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999, 189: 12–19
6. Lorincz AT, Reid R, Jenson AB, Greenberg MD, Lancaster W, Kurman RJ. Human papillomavirus infection of the cervix. *Obstet GynEcol* 1992, 79: 328–337
7. Schiffman MH, Bauer HM, Hoover RN, Glass AG, Cadell DM, Rush BB, Scott DR, *et al.* Epidemiologic evidence showing that human papillomavirus infection causes most cervical intraepithelial neoplasia. *J Natl Cancer Institute* 1993, 85: 958–964
8. Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. *Lancet* 2007, 370: 890–907
9. Koliopoulos G, Nyaga VN, Santesso N, Bryant A, Martin-Hirsch PP, Mustafa RA, Schünemann H, *et al.* Cytology versus HPV testing for cervical cancer screening in the general population. *Cochrane Database Systatic Rev* 2017, 8: CD008587
10. Champer M, Wong AM, Champer J, Brito IL, Messer PW, Hou JY, Wright JD. The role of the vaginal microbiome in gynaecological cancer. *BJOG* 2018, 125: 309–315
11. Tamarelle J, Thiébaut AC, de Barbeyrac B, Bébéar C, Ravel J, Delarocque-Astagneau E. The vaginal microbiota and its association with human papillomavirus, *Chlamydia trachomatis*, *Neisseria gonorrhoeae* and *Mycoplasma genitalium* infections: a systematic review and meta-analysis. *Clin Microbiol Infect* 2019, 25: 35–47
12. Di Pietro M, Filardo S, Porpora MG, Recine N, Latino MA, Sessa R. HPV/Chlamydia trachomatis co-infection: metagenomic analysis of cervical microbiota in asymptomatic women. *New Microbiol* 2018, 41: 34–41
13. Kyrgiou M, Mitra A, Moscicki AB. Does the vaginal microbiota play a role in the development of cervical cancer? *Transl Res* 2017, 179: 168–182
14. Mitra A, MacIntyre DA, Lee YS, Smith A, Marchesi JR, Lehne B, Bhatia R, *et al.* Cervical intraepithelial neoplasia disease progression is associated with increased vaginal microbiome diversity. *Sci Rep* 2015, 5: 16865
15. Gur C, Ibrahim Y, Isaacson B, Yamin R, Abed J, Gamliel M, Enk J, *et al.* Binding of the Fap2 protein of Fusobacterium nucleatum to human inhibitory receptor TIGIT protects tumors from immune cell attack.

Immunity 2015, 42: 344–355

16. Anahtar MN, Byrne EH, Doherty KE, Bowman BA, Yamamoto HS, Soumillon M, Padavattan N, *et al.* Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. *Immunity* 2015, 42: 965–976

17. Voravuthikunchai SP, Bilasoi S, Supamala O. Antagonistic activity against pathogenic bacteria by human vaginal lactobacilli. *Anaerobe* 2006, 12: 221–226

18. Ahmed KA, Xie Y, Zhang X, Xiang J. Acquired pMHC I complexes greatly enhance CD4 + Th cell's stimulatory effect on CD8 + T cell-mediated diabetes in transgenic RIP-mOVA mice. *Cell Mol Immunol* 2008, 5: 407–415

19. Edelman SM, Lehti TA, Kainulainen V, Antikainen J, Kylväjä R, Baumann M, Westerlund-Wikström B, *et al.* Identification of a high-molecular-mass Lactobacillus epithelium adhesin (LEA) of Lactobacillus crispatus ST1 that binds to stratified squamous epithelium. *Microbiology* 2012, 158: 1713–1722

20. Kumar R, Eipers P, Little RB, Crowley M, Crossman DK, Lefkowitz EJ, Morrow CD. Getting started with microbiome analysis: sample acquisition to bioinformatics. *Curr Protocols Hum Genet* 2014, 82: 18.8.1–29

21. Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 2011, 27: 2957–2963

22. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016, 13: 581–583

23. Estaki M, Jiang L, Bokulich NA, McDonald D, González A, Kosciolek T, Martino C, *et al.* QIIME 2 enables comprehensive end-to-end analysis of diverse microbiome data and comparative studies with publicly available data. *Curr Protoc Bioinformatics* 2020, 70: e100

24. Dhariwal A, Chong J, Habib S, King IL, Agellon LB, Xia J. MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res* 2017, 45: W180–W188

25. Forsberg EM, Huan T, Rinehart D, Benton HP, Warth B, Hilmers B, Siuzdak G. Data processing, multi-omic pathway mapping, and metabolite activity analysis using XCMS Online. *Nat Protoc* 2018, 13: 633–651

26. Guijas C, Montenegro-Burke JR, Domingo-Almenara X, Palermo A, Warth B, Hermann G, Koellensperger G, *et al.* METLIN: a technology platform for identifying knowns and unknowns. *Anal Chem* 2018, 90: 3156–3164

27. Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G, Wishart DS, *et al.* MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res* 2018, 46: W486–W494

28. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res* 2021, 49: D545–D551

29. Asnicar F, Weingart G, Tickle TL, Huttenhower C, Segata N. Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *Peer J* 2015, 3: e1029

30. Nejman D, Livyatan I, Fuks G, Gavert N, Zwang Y, Geller LT, Rotter-Maskowitz A, *et al.* The human tumor microbiome is composed of tumor type–specific intracellular bacteria. *Science* 2020, 368: 973–980

31. Klein C, Gonzalez D, Samwel K, Kahesa C, Mwaiselage J, Aluthge N, Fernando S, *et al.* Relationship between the cervical microbiome, HIV status, and precancerous lesions. *mBio* 2019, 10: e02785-18

32. Godoy-Vitorino F, Romaguera J, Zhao C, Vargas-Robles D, Ortiz-Morales G, Vázquez-Sánchez F, Sanchez-Vázquez M, *et al.* Cervicovaginal fungi and bacteria associated with cervical intraepithelial neoplasia and high-risk human papillomavirus infections in a hispanic population. *Front Microbiol* 2018, 9: 2533

33. Paavonen J, Brunham RC. Bacterial vaginosis and desquamative inflammatory vaginitis. *N Engl J Med* 2018, 379: 2246–2254

34. Mitra A, MacIntyre DA, Ntritsos G, Smith A, Tsilidis KK, Marchesi JR, Bennett PR, *et al.* The vaginal microbiota associates with the regression of untreated cervical intraepithelial neoplasia 2 lesions. *Nat Commun* 2020, 11: 1999

35. Usyk M, Zolnik CP, Castle PE, Porras C, Herrero R, Gradissimo A, Gonzalez P, *et al.* Cervicovaginal microbiome and natural history of HPV in a longitudinal study. *PLoS Pathog* 2020, 16: e1008376

36. Mastromarino P, Di Pietro M, Schiavoni G, Nardis C, Gentile M, Sessa R. Effects of vaginal lactobacilli in Chlamydia trachomatis infection. *Int J Med Microbiol* 2014, 304: 654–661

37. Graver MA, Wade JJ. The role of acidification in the inhibition of Neisseria gonorrhoeae by vaginal lactobacilli during anaerobic growth. *Ann Clin Microbiol Antimicrob* 2011, 10: 8

38. Stoyancheva G, Marzotto M, Dellaglio F, Torriani S. Bacteriocin production and gene sequencing analysis from vaginal Lactobacillus strains. *Arch Microbiol* 2014, 196: 645–653

39. Kabuki T, Saito T, Kawai Y, Uemura J, Itoh T. Production, purification and characterization of reutericin 6, a bacteriocin with lytic activity produced by Lactobacillus reuteri LA6. *Int J Food Microbiol* 1997, 34: 145–156

40. Borgdorff H, Gautam R, Armstrong SD, Xia D, Ndayisaba GF, van Teijlingen NH, Geijtenbeek TBH, *et al.* Cervicovaginal microbiome dysbiosis is associated with proteome changes related to alterations of the cervicovaginal mucosal barrier. *Mucosal Immunol* 2016, 9: 621–633

41. Ley RE. Prevotella in the gut: choose carefully. *Nat Rev Gastroenterol Hepatol* 2016, 13: 69–70

42. Jia YJ, Liao Y, He YQ, Zheng MQ, Tong XT, Xue WQ, Zhang JB, *et al.* Association between oral microbiota and cigarette smoking in the Chinese population. *Front Cell Infect Microbiol* 2021, 11: 658203

43. Randis TM, Ratner AJ. Gardnerella and Prevotella: co-conspirators in the pathogenesis of bacterial vaginosis. *J Infect Dis* 2019, 220: 1085–1088

44. Si J, You HJ, Yu J, Sung J, Ko GP. Prevotella as a hub for vaginal microbiota under the influence of host genetics and their association with obesity. *Cell Host Microbe* 2017, 21: 97–105

45. Yang K, Xia B, Wang W, Cheng J, Yin M, Xie H, Li J, *et al.* A comprehensive analysis of metabolomics and transcriptomics in cervical cancer. *Sci Rep* 2017, 7: 43353

46. Hou Y, Yin M, Sun F, Zhang T, Zhou X, Li H, Zheng J, *et al.* A metabolomics approach for predicting the response to neoadjuvant chemotherapy in cervical cancer patients. *Mol Biosyst* 2014, 10: 2126–2133

47. Corn KC, Windham MKA, Rafat M. Lipids in the tumor microenvironment: from cancer progression to treatment. *Prog Lipid Res* 2020, 80: 101055

48. Säde E, Johansson P, Heinonen T, Hultman J, Björkroth J. Growth and metabolic characteristics of fastidious meat-derived Lactobacillus algidus strains. *Int J Food Microbiol* 2020, 313: 108379

49. Wolken WAM, Lucas PM, Lonvaud-Funel A, Lolkema JS. The mechanism of the tyrosine transporter TyrP supports a proton motive tyrosine decarboxylation pathway in *Lactobacillus brevis*. *J Bacteriol* 2006, 188: 2198–2206

50. Jiang M, Xu G, Ni J, Zhang K, Dong J, Han R, Ni Y. Improving soluble expression of tyrosine decarboxylase from *Lactobacillus brevis* for tyramine synthesis with high total turnover number. *Appl Biochem Biotechnol* 2019, 188: 436–449

51. Wu CH, Ko JL, Liao JM, Huang SS, Lin MY, Lee LH, Chang LY, *et al.* D-methionine alleviates cisplatin-induced mucositis by restoring the gut microbiota structure and improving intestinal inflammation. *Ther Adv Med Oncol* 2019, 11: 1758835918821021