HPV-associated cervicovaginal microbiome and host metabolome characteristics

Yao Zhang[2,3,4†], Xu Wu[3,4†], Dan Li[2†], Rong Huang[2], Xiangyu Deng[2], Mingxing Li[3,4], Fukuan Du[3,4], Yueshui Zhao[3,4], Jing Shen[3,4], Yu Chen[3,4], Pingxiu Zhang[5], Congcui Hu[6], Zhangang Xiao[3,4*] and Qinglian Wen[1,2*]

Abstract

Background Cervicovaginal microbiome plays an important role in the persistence of HPV infection and subsequent disease development. However, cervicovaginal microbiota varied cross populations with different habits and regions. Identification of population-specific biomarkers from cervicovaginal microbiota and host metabolome axis may support early detection or surveillance of HPV-induced cervical disease at all sites. Therefore, in the present study, to identify HPV-specific biomarkers, cervicovaginal secretion and serum samples from HPV-infected patients (HPV group, $n = 25$) and normal controls (normal group, $n = 17$) in Xichang, China were collected for microbiome (16S rRNA gene sequencing) and metabolome (UHPLC-MS/MS) analysis, respectively.

Results The results showed that key altered metabolites of 9,10-DiHOME, *α*-linolenic acid, ethylparaben, glycocholic acid, pipecolic acid, and 9,12,13-trihydroxy-10(E),15(Z)-octadecadienoic acid, correlating with *Sneathia (Sneathia_ amnii), Lactobacillus (Lactobacillus_iners), Atopobium, Mycoplasma*, and *Gardnerella*, may be potential biomarkers of HPV infection.

Conclusion The results of current study would help to reveal the association of changes in cervicovaginal microbiota and serum metabolome with HPV infections.

Keywords Human papillomavirus, Cervical screening, Cervicovaginal microbiota, Metabolomics

†Yao Zhang, Xu Wu and Dan Li have contributed equally to this work.

*Correspondence:
Zhangang Xiao
zhangangxiao@swmu.edu.cn
Qinglian Wen
wql73115@163.com
[1]Department of Radiation Oncology, Cancer Center, West China Hospital, Sichuan University, Chengdu, Sichuan, China
[2]Department of Oncology, The Affiliated Hospital, Southwest Medical University, Luzhou, Sichuan, China
[3]Cell Therapy & Cell Drugs Key Laboratory of Luzhou, Department of Pharmacology, School of Pharmacy, Southwest Medical University, Luzhou, Sichuan, China
[4]South Sichuan Institute of Translational Medicine, Luzhou, Sichuan, China
[5]Yanyuan County Maternal and Child Health and Family Planning Service Center, Xichang, Sichuan, China
[6]Yanyuan County People's Hospital, Xichang, Sichuan, China

Background

Cervical cancer is one of the leading causes of cancer death in women [1]. It was reported in 2020 that the age-standardized global cervical cancer incidence was 13.3 cases per 100,000 women and the mortality rate was 7.2 deaths per 100,000 women [2]. In China, the age-standardized incidence rate in 2022 was 13.83 per 100,000, and the mortality rate was 4.54 per 100,000 [3]. Almost all cervical cancer (>95%) are caused by persistent infections with one of the high-risk human papilloma virus (HPV) genotypes [4]. HPV is the most common sexually transmitted infection (STI) [5]. Risk factors include number of sextual partners, age at first sexual intercourse, immunosuppression (including HIV infection and immunosuppressive medications), altered vaginal microbiota

and the presence of other STIs, such as herpes simplex [5, 6]. Some studies pointed out the relationship of HPV with age, gender, smoking, alcohol, and use of oral contraceptives [7]. Because of the vast public health burden of HPV-related cervical cancer in women, there is a necessity for identifying new secondary preventive interventions. Serum-based and/or non-serum-based biomarkers display an ideal modality for early detection or surveillance of HPV-induced cervical cancer at all sites.

There is growing evidence that the cervicovaginal microbiome plays an important role in the persistence of HPV infection and subsequent disease development [8–10]. Studies have shown that women with HPV infection had low *lactobacilli* abundance and high microbial diversity in their cervicovaginas [11], which were positively correlated with the severity of cervical lesions and specific HPVs [12]. It is highlighted that cervical microbiome has a great influence on the outcome of HPV infections [13]. Notably, HPV infection and the associated cervicovaginal microbiome varied significantly cross populations with different habits and regions [14]. Thus, specifically altered cervical bacteria may be potential biomarkers of HPV infection in different populations. Furthermore, previous study indicated that serum circulating HPV DNAs, as well as cytokine levels, vitamins and cofactors, and various gene polymorphisms have been explored [15]. A recent study showed that HPV infection remarkably changed vaginal metabolome regarding the biogenic amines, glutathione, and lipid metabolites [16]. Chorna et al. revealed an association of cervicovaginal microbiome and urine metabolome [17]. However, few studies have focused on serum metabolome. HPV infection may induce characteristic signature change of host small-molecular metabolites. Therefore, the aim of present study is to identify characteristic cervical microbiome and host metabolome for novel HPV-associated biomarkers.

It is acknowledged that, in resource-poor regions where show high-risk factors such as high fertility, early pregnancy, and early marriage, women are at a high risk of HPV infections. Xichang city is one of resource-poor areas in southwestern China, with a population of more than 4.8 million. In this study, the HPV-associated alteration of cervical microbiome and host metabolome in Xichang, China was identified through 16S ribosomal RNA gene (16S rRNA) sequencing and untargeted metabolomics. The results would help to reveal the association of altered cervical microbes and serum metabolites with the HPV infection.

Materials and methods
Subjects
Subjects who volunteered to participate in biomarker study were recruited from a cross-sectional study from 2019 to 2020 at Yanyuan County People's Hospital and Yanyuan County Maternal and Child Health and Family Planning Service Center in Xichang City, China. All subjects ($n=76$) underwent cervical screening via professional gynecological examination, thinprep cytology testing (TCT), HPV test, colposcopy, and biopsy according to the 2019 ASCCP management consensus guidelines [18]. During cervical screening, those who had menstruation, vaginal medication in the last week, post-total hysterectomy, and pregnancy were excluded for the test. The related information of all subjects (BMI, age, history of miscarriage, history of pregnancy, contraceptive method, history of menopause, etc.) was recorded. Women with a recent history of antibiotic use (less than 4 week), anti-infective treatment, urinary incontinence, hysterectomy, and pregnancy were excluded for biomarker study. All women were not given any HPV-related vaccines before. Initially included subjects underwent cytological test by TCT, and patients with TCT=ASC-US/HPV+, or TCT>ASC-US results were subjected for colposcopy. Cervical biopsy followed by histological examination (H&E stain) was conducted if suspicious lesions were identified under colposcopy. The results of the cervical biopsy were graded as normal, chronic cervicitis, CIN1, CIN2/3 and invasive cancer. Finally, all subjects were grouped into: Control group (HPV(-)/TCT(-)) and HPV group (HPV(+)/TCT(+)). Particularly, the HPV group contained those with chronic cervicitis ($n=7$), CIN1 ($n=8$), or CIN2/3 ($n=10$). All subjects gave their informed consent for inclusion before they participated in the study. The study was performed in accordance with the Declaration of Helsinki, and was approved by the Clinical Trial Ethics Committee of Southwest Medical University Hospital (approval number: KY2019039).

Cervicovaginal secretion samples were collected from the peripheral cervical area and cervical secretions, scraped with a cervical swab, placed in cryopreservation tubes and stored in liquid nitrogen. Venous blood was collected in 2-mL sterilized tube and serum samples were obtained by centrifugation at 4℃ (3000 rpm/min, 5 min). The serum was stored in liquid nitrogen before use. All collected samples were further processed and analyzed within 1 month.

Microbiome analysis
Analysis of cervicovaginal microbiome was performed based on the short-read 16S rRNA sequencing [19, 20]. This technique offers high-throughput examination of microbial changes, which is convenient and powerful. However, application of 16S gene requires some assumptions, e.g., sequences of >95% similarity represent the same genus, and sequences of >97% similarity represent the same species [21]. Therefore, the resolution of 16S

rRNA sequencing at species level is generally considered low [21].

DNA extraction

Total genome DNA from cervicovaginal secretion samples was extracted using the cetyltrimethylammonium bromide (CTAB) method [22]. DNA concentration and purity was monitored on 1% agarose gels. DNA was diluted to 1 ng/μL using sterile water.

High-throughput 16S rRNA sequencing

The amplification was conducted using the primers (341 F: 5-CCTAYGGGRBGCASCAG-3; 806R: 5-GGAC-TACNNGGGTATCTAAT-3) targeting the V3 and V4 hypervariable regions of the 16S rRNA gene. To differentiate each sequenced specimen and acquire accurate phylogenetic and taxonomic data, the gene products were obtained with forward and reverse error-correcting barcodes. The amplified PCR products were separated by electrophoresis on 2% agarose gel and purified by the GeneJETTM Gel Extraction Kit (Thermo Scientific, USA). Sequencing libraries were obtained using Ion Plus Fragment Library Kit 48 Rxns (Thermo Scientific) following manufacturer's instructions. The quality of library was evaluated on the Qubit@ 2.0 Fluorometer (Thermo Scientific). Eventually, the library was sequenced on an Ion S5TM XL platform [23].

Data analysis

Quality filtering on the raw reads were conducted to acquire the high-quality clean reads based on the Cutadapt (V1.9.1) quality control protocol. Chimera sequences, which were detected by comparison of the reads with reference database (Silva database) by UCHIME algorithm, were removed. Sequence analysis was conducted using Uparse software (Uparse v7.0.1001). Sequences with high similarity (≥97%) were assigned to the same operational taxonomic units (OTUs). Representative sequence for each OTU was used for annotation of taxonomic information against Silva Database. MUSCLE software (Version 3.8.31) was used to study phylogenetic relationship of varied OTUs, and the difference of the dominant species in different groups. Abundance of OTU was normalized based on a standard of sequence number corresponding to the sample with the least sequences.

Alpha diversity indices (Observed species, Chao1, Shannon, Simpson, and ace) were calculated by QIIME software (Version 1.9.1). PCoA analysis was performed based on R software (Version 2.15.3). Linear discriminant analysis (LDA) with effect size (LEfSe) was applied to assess the differentially abundant taxon.

Metabolomics analysis
Sample preparation

Serum sample (100 μL) and ice-cold methanol (400 μL) were vortex mixed. The samples were incubated at 4°C for 5 min, followed by centrifugation at 15,000 rpm for 5 min. An aliquot of supernatant was diluted by LC-MS grade water, which were subsequently centrifuged at 15,000 g at 4°C for 10 min. Finally, the supernatant were used for subsequent analysis. QC samples were prepared by mixing equal volume of all samples. For the blank sample, its pretreatment process was the same as experimental sample.

UHPLC-MS/MS analysis

The UHLC-MS/MS analysis was conducted using a Vanquish UHPLC system (Thermo Fisher) with an Orbitrap Q Exactive series mass spectrometer (Thermo Fisher). Samples were injected into an Hyperil Gold column (100×2.1 mm, 1.9 μm) using a 16-min linear gradient at a flow rate of 0.2 mL/min. The mobile phases for the positive ion mode were solvent A (0.1% formic acid in water) and solvent B (Methanol). The mobile phases for the negative ion mode were solvent A (5 mM ammonium acetate, pH 9.0) and B (Methanol). The elution gradient was as follows: 2% B, 1.5 min; 2-100% B, 1.5–12.0 min; 100% B, 12.0–14.0 min; 100-2% B, 14.0–14.1 min; 2% B, 14.1–17 min. The MS parameters in both positive/negative polarity mode were set as: spray voltage, 3.2 kV; capillary temperature, 320°C; sheath gas flow rate, 35 arb; aux gas flow rate, 10 arb.

Data analysis

The raw UHPLC-MS/MS data files were processed by the Compound Discoverer 3.1 (CD3.1, Thermo Fisher) where peak alignment, peak picking, and quantitation for each metabolite were conducted as we previously reported [24]. Normalization of peak intensities were performed according to the total spectral intensity. The normalized data was applied to calculate the molecular formula according to additive ions, molecular ions and fragments. The databases of mzCloud (https://www.mzcloud.org/), mzVault and MassList were used for acquiring the accurate results. Statistical analysis was conducted using the R software (R version R-3.4.3), Python (Python 2.7.6 version) and CentOS (CentOS release 6.6).

Identified metabolites were annotated by the KEGG, HMDB and Lipidmaps databases. The univariate analysis (t-test) was used to calculate p-value. Significantly differential metabolites were identified based on the following rule: VIP>1 and p-value<0.05 and fold change≥1.2 or FC≤0.833. Volcano plot was applied to filter metabolites of interest based on the Log2 (FC) and -log10 (p-value) of each metabolite. Partial least squares discriminant analysis (PLS-DA) plot was performed using SIMCA 13.0.

The functions of differential metabolites were investigated using the KEGG database. The metabolic pathway enrichment of differential metabolites was performed using Metaboanalyst (https://www.metaboanalyst.ca/). The abundance of metabolites was normalized by Autoscaling. Significantly enriched pathways were indicated with adjusted p value<0.05 and at least 2 annotated metabolites. Biomarker models in Metaboanalyst were applied according to the PLS-DA multivariate algorithm to identify specific metabolites. The area under the receiver operating characteristic (ROC) curve was generated and the area under the ROC curve (AUC) was calculated. The cut-off of AUC for candidate biomarker was set at 0.6.

Statistical analysis

Statistical analysis was performed with GraphPad Prism 6.0 (U.S.), and significant differences in clinical characteristics were assessed with unpaired t-test or Fisher's exact test. Pearson correlation analysis was used to assess the relationship between metabolites or between species and metabolites. All results were considered statistically significant at $p<0.05$.

Results

Seventy-six participants were initially included in the microbiome metabolomics study, with 23 of them excluded from the microbiome analysis (12 sampling failures, 9 non-detects and 2 outliers) and 11 excluded from the metabolome analysis due to abnormal values. Paired microbiome and metabolomics samples from the remaining 42 participants (17 in the normal group and 25 in the HPV group) were subjected to microbiome and metabolome analyses. All HPV(+) subjects had infected one of 14 high-risk HPV genotypes: 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66 and 68 (Supplementary Table 1). Seven of them had mixed infection of both low-risk and high-risk HPV (Supplementary Table 1). Analysis of clinical data from the 42 participants showed no statistical differences between the normal and HPV groups prior to collection (Supplementary Table 1), indicating that no knowable influencing factors affecting the analysis of the multi-omics study between the two groups were identified.

Cervicovaginal microbiome structure analysis

After clustering the sequences into OTUs with 97% consistency by default, 625 OTUs were obtained (Supplementary Table 2). To assess the microbial structures, the alpha diversity (Observed species, Shannon, Simpson, Chao1 and Ace, Fig. 1A and E) and beta diversity (PCoA analysis; Fig. 1F and G) of the cervicovaginal microbiota was analyzed cross groups. Although we observed a slight increase of Shannon and Simpson indices in HPV

group compared to the normal group, they together with other indices (Observed species, Chao1 and ACE) did not show statistical differences. Therefore, the alpha diversity was not statistically significant between the two groups. Based on the unweighted PCoA (Fig. 1F) and weighted PCoA (Fig. 1G) plots at OTU level, the bacterial structures in the two groups had some differences yet mostly were similar. The HPV group and normal group samples were not clearly separated (Fig. 1). These results suggest that HPV infection did not significantly alter the alpha diversity of cervicovaginal microbiota and may have an impact on specific microbial structures.

Changes in the composition of the cervicovaginal microbiota associated with HPV infection

Analysis of the top 10 species at the phylum level in the HPV and normal groups revealed significant differences in the cervicovaginal bacteria (Fig. 2). Firmicutes was the most dominant phylum, accounting for 36.07% of the HPV group and 28.03% of the normal group, respectively (Fig. 2A). The HPV group had higher levels of Actinobacteriota (18.17% vs. 2.84%), Fusobacteriota (5.70% vs. 0.26%), Cyanobacteria (0.21% vs. 0.26%) and Caldatribacteriota (0.038% vs. 0.004%) than the normal group. It is indicated that HPV remarkably impacts on cervicovaginal microbial abundance.

Welch's t-test and Mann-Whitney test were performed on the phylum, family, genus and species level to compare the differences in cervicovaginal microbiota between the normal and HPV groups. At the phylum level, the Firmicutes (Fig. 2B) ($p=0.0134$) was higher in the normal group than in the HPV group. The HPV group had higher numbers of the Actinobacteriota (Fig. 2C) ($p=0.0194$), Fusobacteriota (Fig. 2D) ($p=0.0263$), and Cyanobacteria (Fig. 2E) ($p=0.0231$) were also significantly enriched compared to the normal group. Although the upregulation of fecal Firmicutes/Bacteroidetes ratio is considered an indicator of several pathological conditions [16], our results were the opposite (12.37% vs. 15.51%) in cervicovaginal condition. At the family level (Fig. 3A), *Lactobacillaceae* (Fig. 3B) ($p=0.0172$) was more prevalent in the normal group than in the HPV group. Compared to the normal group, *Bifidobacteriaceae* (Fig. 3C) ($p=0.0194$) *Atopobiaceae* (Fig. 3D) ($p=0.0466$), *Mycoplasmaceae* (Fig. 3E) ($p=0.0014$), *Leptotrichiaceae* (Fig. 3F) ($p=0.0254$), *Aerococcaceae* (Fig. 3G) ($p=0.0169$) and *Erysipelotrichaceae* (Fig. 3H) ($p=0.0462$) were enriched in the HPV group.

At the genus level (Fig. 4A), compared to the normal group, the HPV group had significantly higher levels of *Ureaplasma* (Fig. 4B) ($p=0.0106$), *Aerococcus* (Fig. 4C) ($p=0.0105$), *Sneathia* (Fig. 4E) ($p=0.0253$), *Gardnerella* (Fig. 4F) ($p=0.0050$), *Mycoplasma* (Fig. 4G) ($p=0.0021$). There were differences in *Ralstonia* between
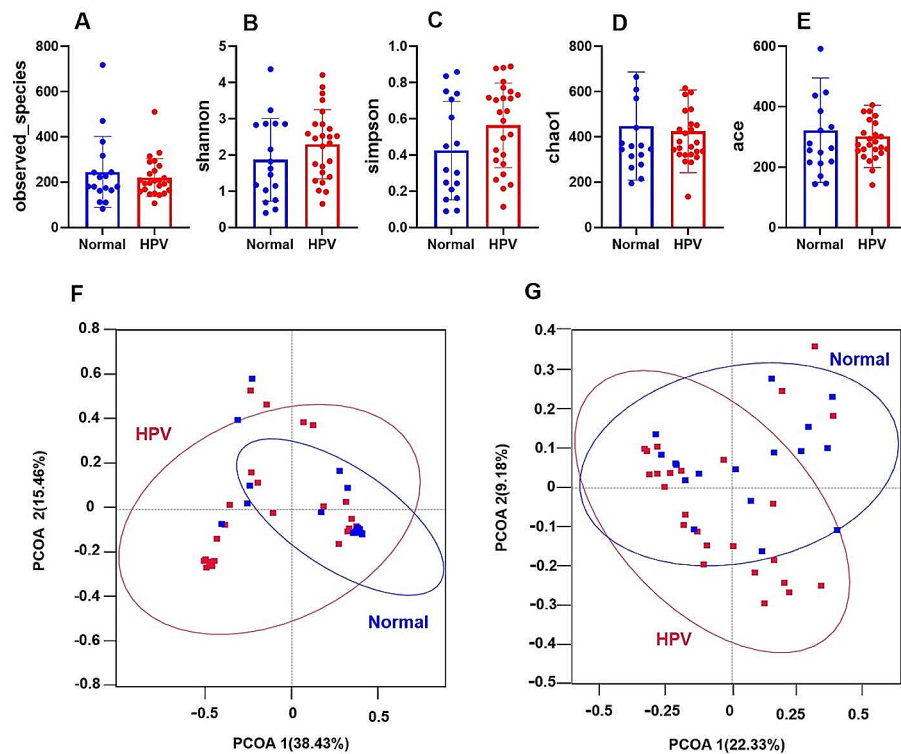
Fig. 1 Analysis of microbial diversity in cervix and vagina. **A–E** Species diversity difference between normal group and HPV group was estimated by Observed-species, Shannon, Simpson, Chao1 and Ace indices. HPV ($n = 25$), group with HPV infected patients; Normal ($n = 17$), subjects with normal cervical condition. **F–G** unweighted and weighted PCoA plot based on OTU levels showing bacterial structure clustering. HPV group (red), Normal group (blue)
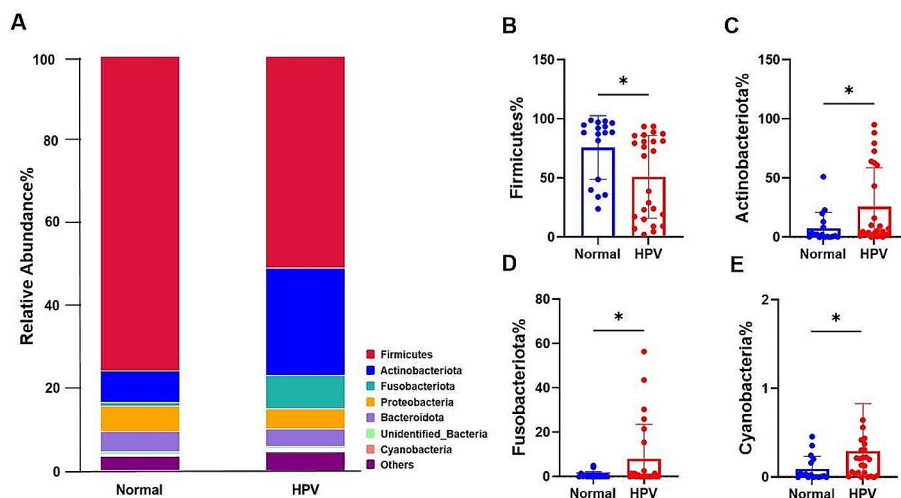

Fig. 2 **A** Relative abundance of the most abundant 10 phylum in the HPV group and Normal group. **B–E** statistically comparison of Firmicutes, Actinobacteriota, Fusobacteriota and Cyanobacteria. *$p < 0.05$

the two groups, but they were not statistically significant (Fig. 4D). The normal group had significantly more *Lactobacillus* (Fig. 4H) ($p = 0.0172$) than the HPV group. At the species level (Fig. 5A), *Lactobacillus_iners* (Fig. 5B) ($p = 0.0196$) in the normal group was also significantly higher than that in the HPV group. Compared to the normal group, the HPV groups had higher abundance

of *Veillonella_montpellie* (Fig. 5C) ($p = 0.0186$), *Ureaplasma_parvum* (Fig. 5D) ($p = 0.0349$), *Sneathia_amnii* (Fig. 5E) ($p = 0.0480$), and *Aerococcus_christensenii* (Fig. 5F) ($p = 0.0096$).

The LEfSe model was then used to identify specific microbiota that may be significantly associated with HPV infection (Fig. 6). The following bacteria were significantly
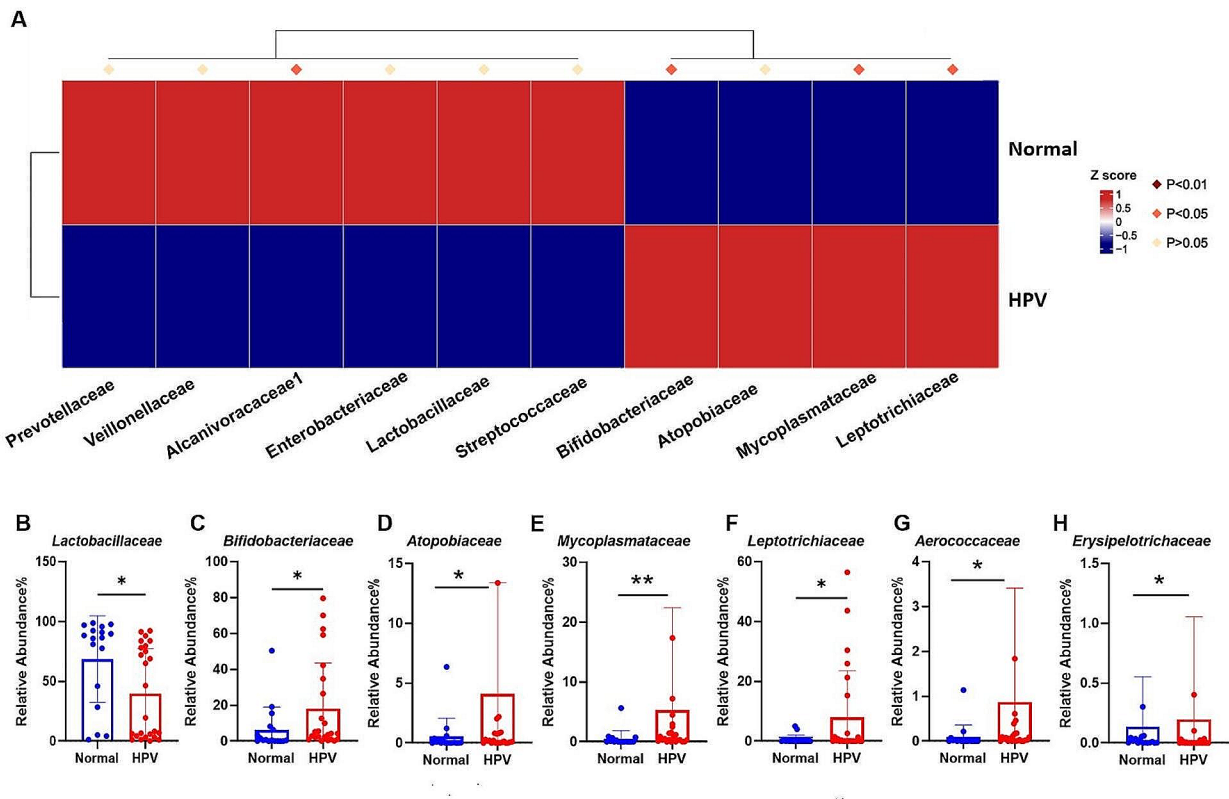
Fig. 3 **A** Comparison of bacterial difference at family level. **B**–**H** the relative abundance of Lactobacillaceae, Bifidobacteriaceae, Atopobiaceae, Mycoplasmaceae, Leptotrichiaceae, Aerococcaceae and Erysipelotrichaceae. *$p < 0.05$, **$p < 0.01$
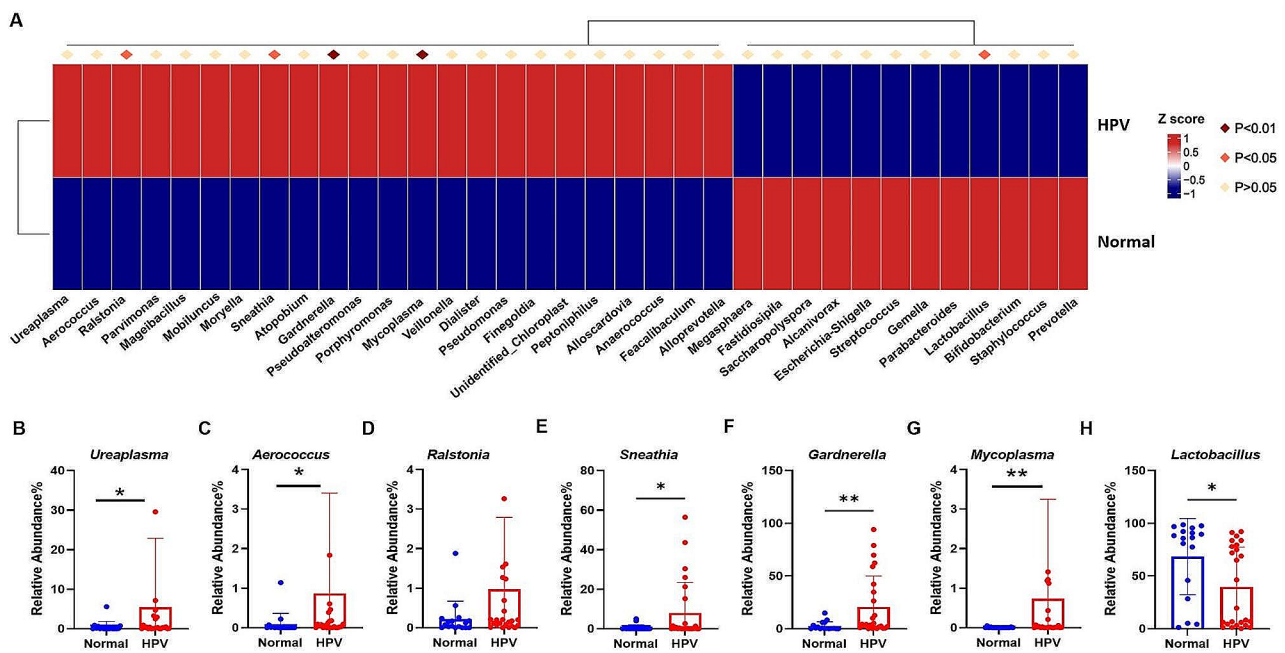

Fig. 4 **A** Comparison of bacterial difference at genus level. **B**–**H** the relative abundance of *Ureaplasma, Aerococcus, Ralstonia, Sneathia, Gardnerella, Mycoplasma, Lactobacillus*. *$p < 0.05$, **$p < 0.01$
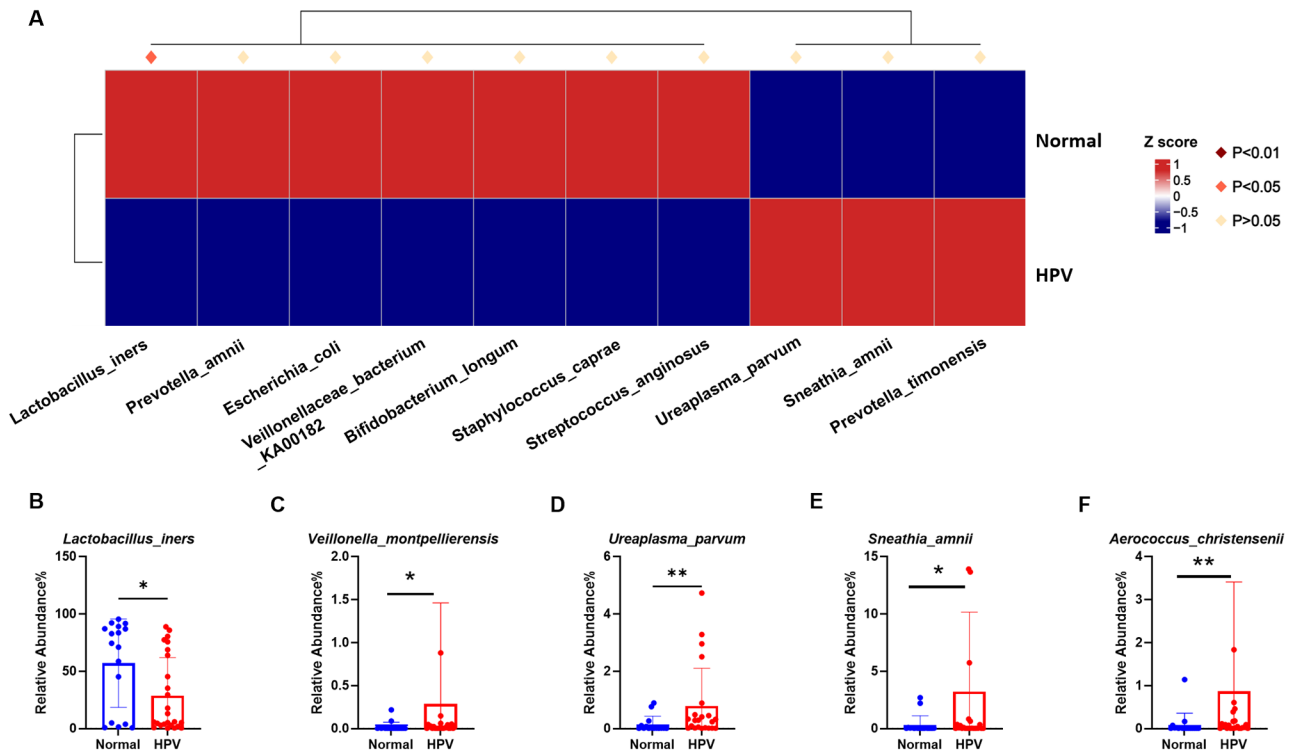
Fig. 5 **A** Comparison of bacterial difference at species level. **B**–**H** the relative abundance of *Lactobacillus_iners, Veillonella_montpellie, Ureaplasma_parvum, Sneathia_amnii* and *Aerococcus_christensenii. *p < 0.05, **p < 0.01
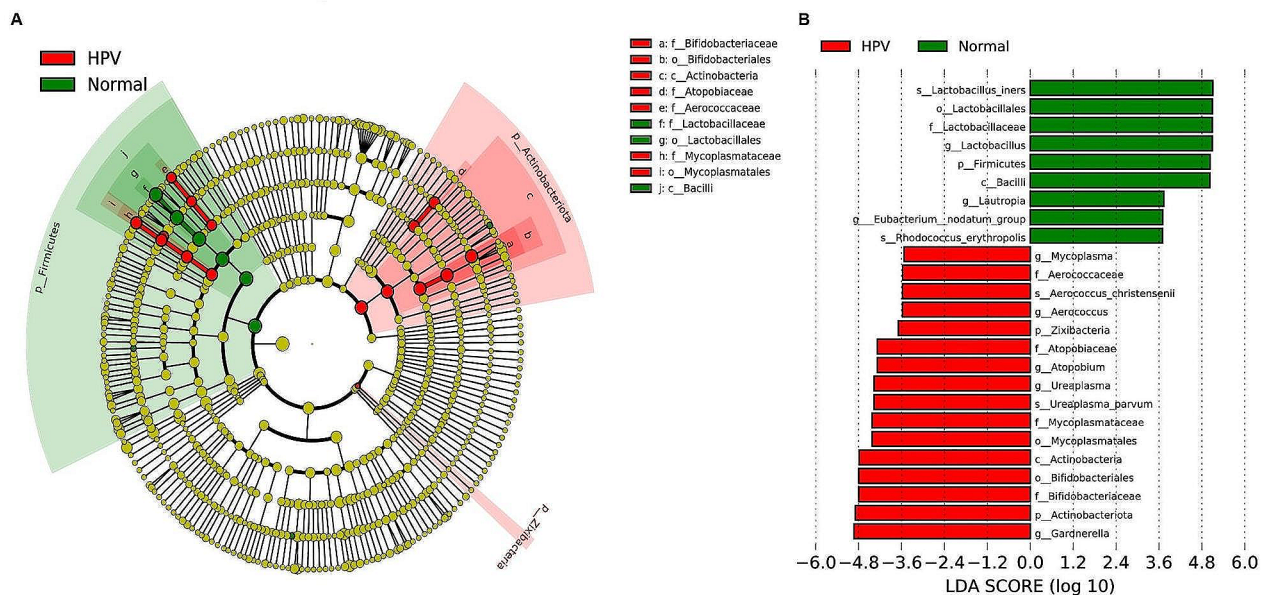

Fig. 6 Linear discriminant analysis (LDA) integrated with effect size (LEfSe) analysis. **A** Cladogram indicating the phylogenetic distribution of microbiota correlated with the Normal or HPV groups. **B** The differences in abundance between the Normal and HPV groups

enriched in the normal group (LDA score > 4.8): phylum of Firmicutes; family of *Lactobacillaceae*; genus of *Lactobacillus*; and species of *Lactobacillus_iners*. We found that the following cervicovaginal microbiota were enriched in the HPV group (LDA score > 3.6): phyla of

Actinobacteriota and Zixibacteria; family of *Bifidobacteriaceae, Mycoplasmataceae,* and *Atopobiaceae*; genus of *Gardnerella, Ureaplasma* and *Atopobium*; and species of *Ureaplasma_parvum.*

Zhang *et al. BMC Microbiology*   (2024) 24:94

Page 8 of 14

Overall overview of serum metabolites in the normal and HPV groups

A growing number of studies have found that cervicovaginal microbes are closely associated with HPV infection, and the development of cervical cancer [9, 25–27]. We hypothesized that changes in human metabolic pathways may be influenced by HPV. Therefore, we performed a metabolomic analysis of serum samples based on UPLC-MS untargeted metabolomics to identify potential HPV associated or microbial-related host metabolites.

We successfully identified 1488 metabolites, from both positive and negative detection mode, in the normal and HPV groups (Supplementary Table 3). A total of 32 differential metabolites were screened from the normal and HPV groups, of which 21 metabolites were down-regulated and 11 metabolites were up-regulated (Fig. 7A). Based on the relative abundance of these differentially expressed metabolites, PLS-DA was used to assess the metabolic differences between the normal and HPV groups. The results showed a significant difference in the distribution of cervicovaginal metabolites between the normal and HPV groups (Fig. 7B).

Differential metabolites were further used to predict KEGG metabolic pathways alterations (Fig. 7C and Supplementary Table 4). The identified impacted pathways included the linolenic acid metabolism, glycerol phospholipid metabolism, arachidonic acid metabolism, unsaturated fatty acid biosynthesis, lysine degradation, pyruvate metabolism, primary cholecystic acid biosynthesis. Pathway enrichment analysis based on the relative abundance of metabolites showed that the KEGG metabolic pathway of α-linolenic acid (ALA) metabolism was significantly enriched (Fig. 7C) and two metabolites (Phosphatidylcholine, (9Z,12Z,15Z)-Octadecatrienoic acid) were directly involved. Enrichment results for specific pathways are shown in Supplementary Table 5. The main enrichment pathway was related to linolenic acid metabolism. Heat map analysis showed significant differences in metabolic patterns between the normal and HPV groups based on the identified biomarkers and metabolites of the enriched pathways (Fig. 8).

Identification of specific metabolites for HPV infection

To further identify metabolite changes associated with HPV infection, biomarker analysis was performed. The area under the ROC curve for the 13 variable model was 0.823 (CI: 0.689–0.94) (Fig. 9A). The predicted probability of classification for all samples indicated that the normal and HPV groups were well classified (Fig. 9B). Figure 9C shows the 13 most important biomarkers, including 9,12,13-Trihydroxy-10(E),15(Z)-octadecadienoic acid, PC (18:0/22:6), PC (18:3 (6Z,9Z,12Z) /18:1(9Z)), α-linolenic acid, 9,10-DiHOME, PC (18:3 (6Z,9Z,12Z) /P-18:1 (11Z)), pipecolic acid, O-cresol, prostaglandin F3α, ethylparaben, S-lactoylglutathione, glycocholic acid, and 3-methylcrotonylglycine. The ROC curves of these potential candidate biomarkers showed that the AUC of all candidate biomarkers was above 0.656 with a P value less than 0.05, indicating that they might be significantly associated with HPV infection (Supplementary Table 6).

Correlation analysis reveals an overview of the microbial metabolic axis of HPV infection

Based on the microbiome and metabolomics data described above, we subsequently performed Pearson correlation analysis to identify relevant microorganisms and metabolites in the HPV and normal groups.

As shown in Fig. 10, 9,10-diHOME was significantly positively correlated with *Sneathia* ($p=0.0006$). The ALA was positively correlated with *Sneathia-amnii* ($p=0.0412$) and *Sneathia* ($p=0.0008$), but was significantly negatively correlated with *Lactobacillus* and *Lactobacillus_iners*. *Ethylparaben* was positively correlated with *Atopobium* ($p=0.0077$), *Sneathia* ($p=0.0077$) and *Mycoplasma* ($p=0.0069$). Glycocholic acid ($p=0.0069$) was positively correlated with *Atopobium* ($p=0.00002$) and *Ralstonia* ($p=0.00001$). Pipecolic acid was negatively correlated with *Sneathia* ($p=0.0161$) and *Mycoplasma* ($p=0.0026$). *Gardnerella* was positively correlated with 9,12,13-Trihydroxy-10(E),15(Z)-octadecadienoic acid ($p=0.0492$). *Lactobacillus_iners* was positively correlated with PC (18: 3(6Z,9Z,12Z) / 18: 1 (9Z)) ($p=0.0120$). The results showed that 9,10-DiHOME, ALA, ethylparaben, glycocholic acid, pipecolic acid, 9,12,13-trihydroxy-10(E),15(Z)-octadecadienoic acid, PC (18:3(6Z,9Z, 12Z) / 18:1(9Z)) and PC (18:3(6Z,9Z,12Z) / P-18: 1 (11Z)), correlating with *Sneathia (Sneathia-amnii), Lactobacillus* (*Lactobacillus_iners), Atopobium, Mycoplasma*, and *Gardnerella*, may be potential biomarkers of HPV infection.

Discussion

HPV prevalence statistics in recent years has been reported in China, ranging from 8.42 to 21.1% [5, 28] The prevalence of HPV as well as the associated microbiome-metabolome feature in Sichuan, particularly in the poor regions, has not yet been investigated. In the present study, we investigate the HPV-associated cervicovaginal microbiome and host metabolome characteristics in western Sichuan, China.

Cervicovaginal microbial structure has been associated with HPV infection [27], high-risk HPV infection type [26], and cervical abnormalities [29, 30]. For the female reproductive tract, health is usually associated with low microbial diversity and the predominance of only one or a few lactic acid bacteria (*lactobacillus* spp., including *L. gasseri*, *L. crispatus*, *L. jensenii*, and *L. iners*) [31–33]. A decrease in *lactobacillus* and an increase in bacterial
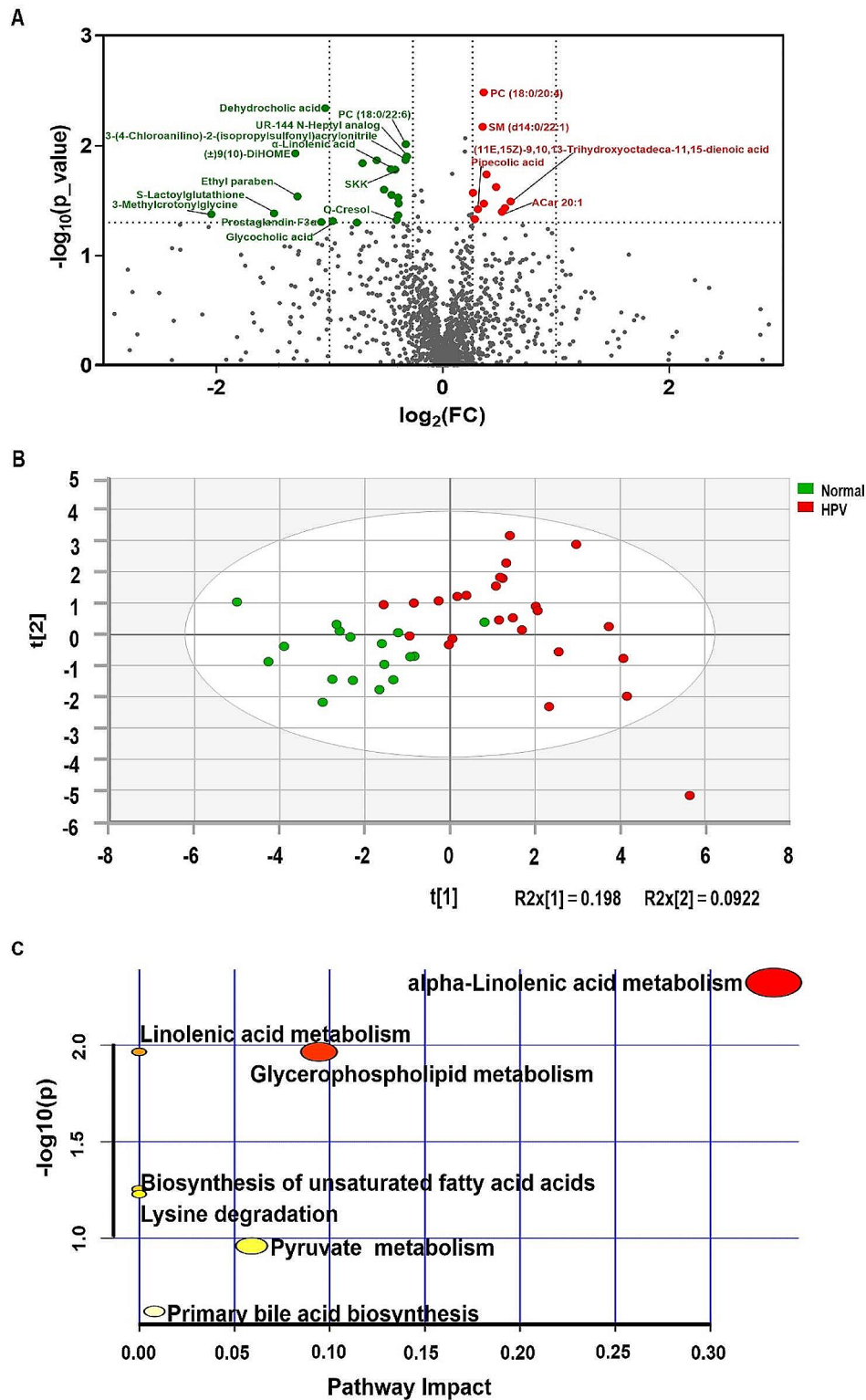
Fig. 7 Serum metabolome analysis. **A** Volcanogram shows differential accumulation of metabolites [log2 (fold-change) on the X-axis] and significant change [-log10 (P) on the Y-axis] between the normal and HPV groups. **B** Partial least squares discriminant analysis (PLS-DA) showed differences between the normal and HPV groups. **C** bubble map showing the enriched metabolic pathway in the normal group and HPV group
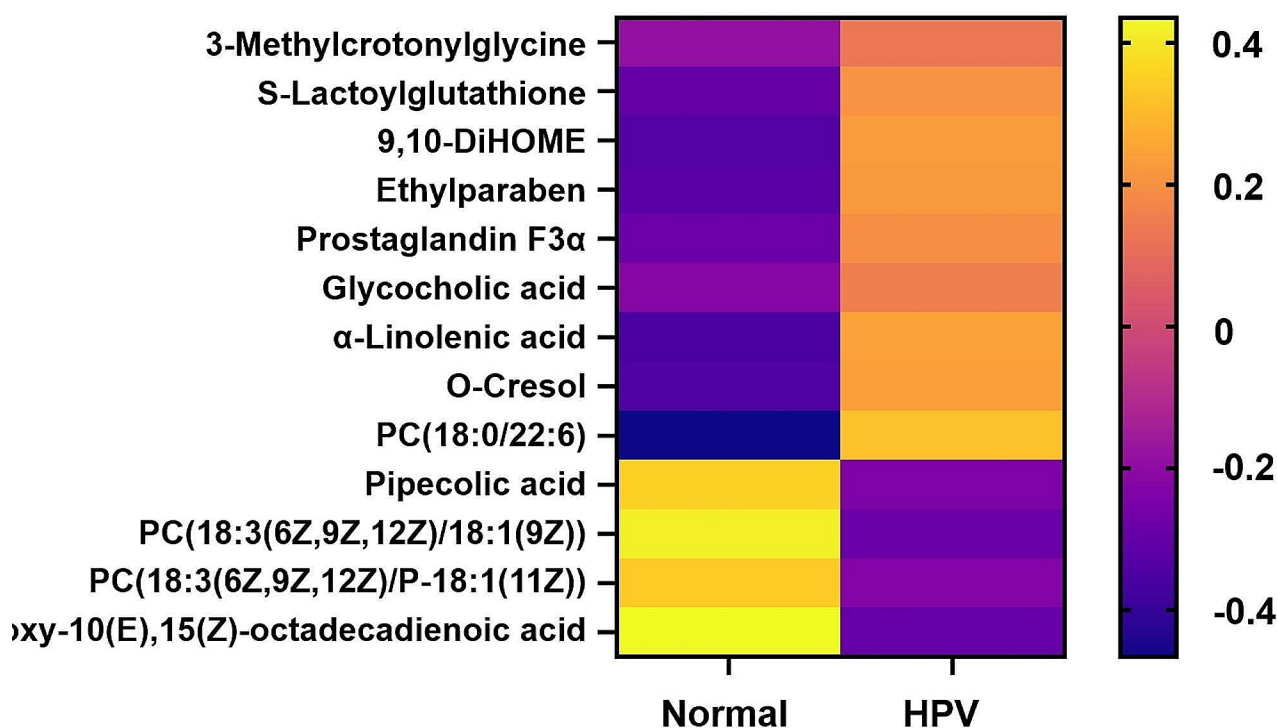
Fig. 8 Heat maps of identified key metabolites for differentiation of normal and HPV groups

species diversity were associated with various cervical diseases such as bacterial vaginosis (BV), vulvovaginal atrophy (VVA) [34], HPV infection [35], and progression of cervical precancerous lesions [12]. Generally, normal cervicovaginal microecology dominated by *lactobacillus* helps to maintain a stable pH at 3.8–4.5, which reduces HPV infection and inhibit persistent HPV infection [36]. The acidic environment constrains the outgrowth of several opportunistically pathogenic strains, such as *Chlamydia trachomatis*, *Gardnerella vaginalis*, and *Neisseria gonorrhoeae* [37]. *Lactobacillus*-derived lactic acid also enhances the viscosity of cervicovaginal mucus, increases viral particle trapping, modulating inflammatory immune response and suppressing the growth and migration of cervical cancer cells [38]. In addition, other beneficial microbiota can also generate antimicrobial peptides of bacteriocins, which exhibits bactericidal effect and regulating inflammation [39]. Consistently, the normal group in the present study was dominated by *Lactobacillus* and the abundance of other microorganisms was low. On the contrary, HPV infection led to significantly reduced *Lactobacillus*, particularly for *Lactobacillus_iners.* Other *Lactobacillus spp.* were not significantly altered.

Studies have shown that *Atopobium* and *sialidase* genes could serve as microbial markers of persistent HPV infection [36]. The finding that *Ureaplasma parvum* infection expressed *UpVF* and induced vacuolated HeLa cell death suggested a potential pathogenic mechanism between

Ureaplasma parvum infection and cervical carcinogenesis [40]. *Sneathia* strains was positively associated with cervical cancer [41, 42]. Particularly, *Sneathia_amnii* was reported associated with bacterial vaginosis, preeclampsia, preterm labor, spontaneous abortion, postpartum bacteremia, and other invasive infections [42]. *Mycoplasma* infection increased the prevalence of high-risk HPV infection, which may be a factor in the persistence of high-risk HPV infection that exacerbated cervical lesions [43]. *Veillonella_montpellierensis* infection may cause disease by depleting cervicovaginal lactate [44]. *Aerococcus_christensenii* caused a potentially life-threatening membrane infection in the vagina and cervix [45]. In the present study, the results showed that HPV infection resulted in increased *Ureaplasma* (*Ureaplasma_parvum*), *Aerococcus* (*Aerococcus_christensenii*), *Sneathia* (*Sneathia_amnii*), *Gardnerella*, *Mycoplasma* and *Veillonella_montpellie*. LefSe analysis showed that the phylum-level Actinobacteriota and the genus-level *Gardnerella*, *Ureaplasma* and *Atopobium* may be potential features of HPV infection, both of which have high LDA scores. These findings were generally consistent with previous documents. It is suggested that there was a complex association and vicious circle among bacterial changes, HPV infection, and cervical lesions, which together contributed to the development and progression of cervical disease initiation and progression.
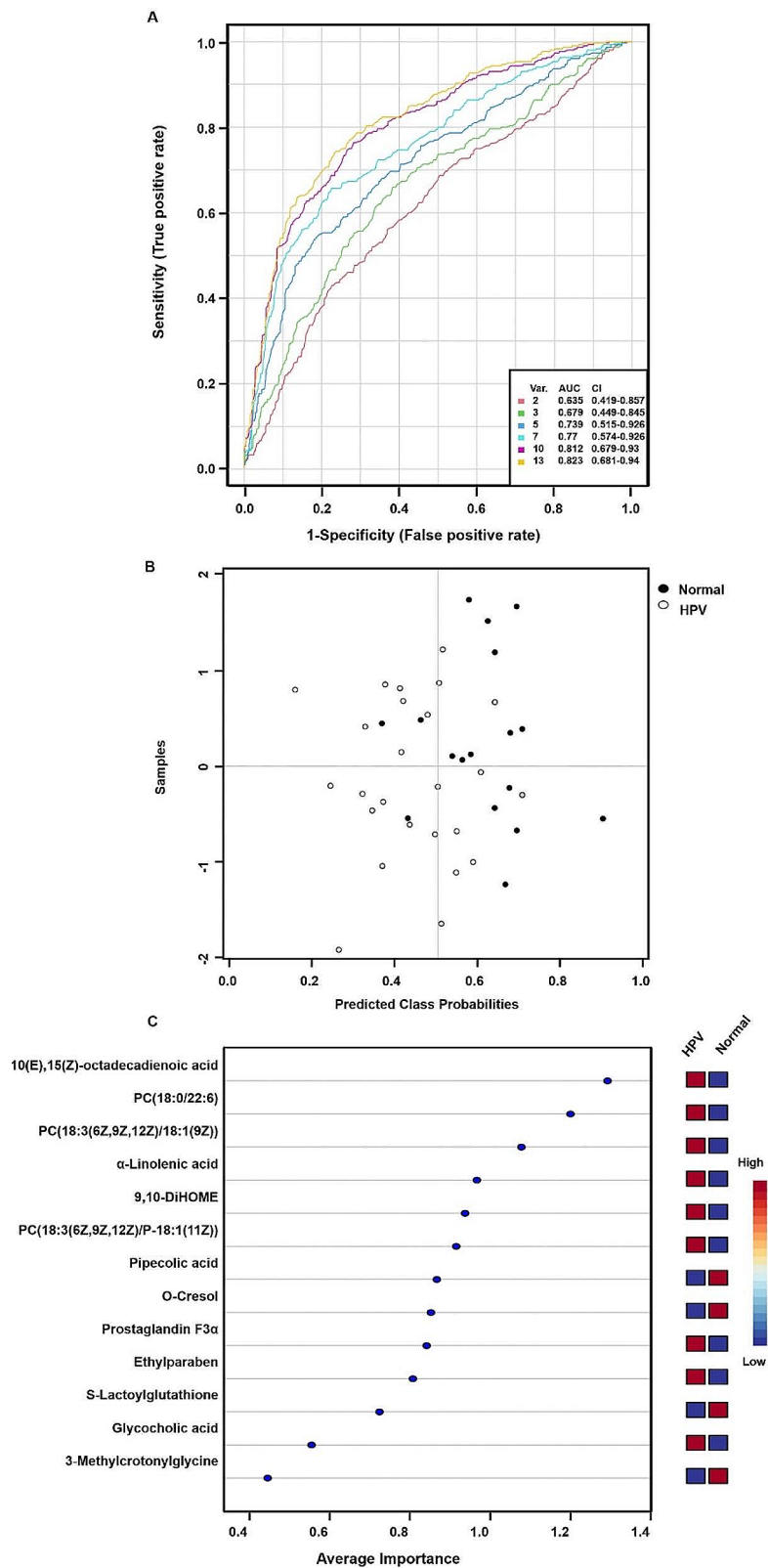
Fig. 9 Important discriminatory metabolites identified by correlation and multivariate analysis between the Normal and HPV groups. **A** ROC curve based on cross validation (CV) performance. The variables used in the model are displayed in **C** below. **B** the predictive class probabilities for each sample based on AUC. **C** The PLS-DA model obtained significant discriminatory metabolites of average importance
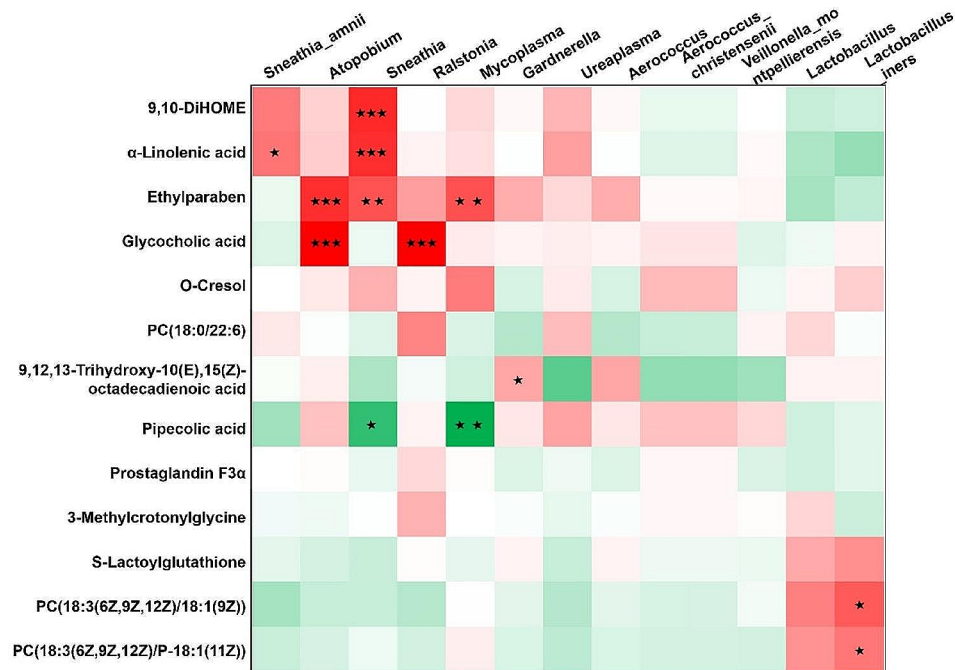
Fig. 10 Integrated correlation-based analysis (Pearson's correlation) of key altered microbes and metabolites upon HPV infection

Subsequently, an untargeted metabolomic analysis of serum samples from HPV-infected patients was performed, cooperating with microbial analysis to identify potential metabolic biomarkers for HPV infection. Through multiple comparison, multivariate and correlation analysis, we identified several significantly altered metabolites, including 9,10-DiHOME, ALA, ethylparaben, glycocholic acid, pipecolic acid, 9,12,13-trihydroxy-10(E),15(Z)-octadecadienoic acid, PC (18:3(6Z,9Z, 12Z) / 18:1(9Z)) and PC (18:3(6Z,9Z,12Z) / P-18: 1 (11Z)). We established a 13-variable model (containing the 13 significantly altered metabolites) to reveal their association with HPV infection. Notably, the area under the ROC curve for the 13-variable model was 0.823 (CI: 0.689–0.94), although the area under the ROC curves for all individual metabolites ($\geq$0.656, $p < 0.05$) were not high. It is thus highlighted that these significantly altered metabolites were potential candidate biomarkers that were significantly associated with HPV infection and cervical bacteria alterations. Previous studies showed that ALA was found to reduce the expression of HPV tumor proteins E6 and E7 and restore the expression of tumor suppressor proteins p53 and Rb [46] and have oncogenic effects. Meanwhile, ALA suppressed cell proliferation in prostate, breast and bladder cancers [47, 48] and up-regulated tumor suppressor genes, playing an suppressive role on various tumors [49]. The correlation between placental ethylparaben and cord blood gamma-glutamate transferase (GGT) and glucose levels provides a starting point for further studies on the mechanisms of uterine metabolic processes associated with parabens

[50]. Glycocholine was used as a metabolic biomarker for traumatic brain injury (TBI) via assessing the duration of TBI injury [51]. Glycocholic acid metabolism was closely associated with certain strains of *Bacteroides* [52]. Either in anaerobic or aerobic cultures *Bacteroides* are unable to convert the side chains of bile acids. ($\pm$)9(10)-DIHOME induced oxidative stress at high concentrations [53]. Through correlation analysis, we found that there was strong correlation among cervical microbiota alteration and serum metabolism during HPV infections. For instance, ALA and *Lactobacillus* were closely related in HPV infection.

Omics study has been widely applied for biomarker screening and identification in the past decade. Disease-specific biomarkers facilitate diagnosis of diseases as well as monitoring of response to therapy in individuals. However, discovering disease-specific biomarkers faces several challenges, such as determining the causal relationships of observed changes, understanding their functions in disease initiation and progression, and the individual or populational heterogeneity [54]. Therefore, the identified potential biomarkers from microbiome and metabolome analysis in the present study still require further validation in larger clinical cohort and future causality study based on the understanding of disease mechanism.

The limitation of current study included the following. Firstly, the sample size of this study was considered small. To propose the identified microbes and/or serum metabolites as biomarkers, the findings require further unbiased validations in an independent and larger

sample cohort. Besides, due to the small sample size, it is unable to see the differences depending on CIN grades, HPV types and vaginal community state types (CSTs). Secondly, the current methodology did not take into consideration that whether HPV persistently or non-persistently infected the included subjects. Some of the CIN subjects with HPV infection had the chance to be recovered to normalcy due to clearance by host immunity. Thirdly, this study did not perform causality experiments to confirm whether the identified metabolic markers were specific to vaginal microbiome alteration.

Conclusion

In this study, we showed that 9,10-DiHOME, ALA, ethylparaben, glycocholic acid, pipecolic acid, 9,12,13-trihydroxy-10(E),15(Z)-octadecadienoic acid, PC (18:3(6Z,9Z, 12Z) / 18:1(9Z)) and PC (18:3(6Z,9Z,12Z) / P-18: 1 (11Z)), correlating with *Sneathia (Sneathia-amnii), Lactobacillus* (*Lactobacillus_iners), Atopobium, Mycoplasma*, and *Gardnerella,* may be potential candidates for biomarker of HPV infection.

Abbreviations
ALA	α-linolenic acid
AUC	area under the curve
CI	confidence interval
GM-CSF	granulocyte-macrophage colony-stimulating factor
GGT	gamma-glutamate transferase
HPV	human papillomavirus
LDA	linear discriminant analysis
PCoA	principal coordinate analysis
ROC	receiver operating characteristic
16S rRNA	16S ribosomal ribonucleic acid
TCT	Thinprep cytology testing
VVA	vulvovaginal atrophy

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12866-024-03244-1.

Supplementary Material 1

Declarations

References
1. Mattiuzzi C, Lippi G. Cancer statistics: a comparison between World Health Organization (WHO) and global burden of Disease (GBD). Eur J Pub Health. 2020;30(5):1026–7.
2. Singh D, et al. Global estimates of incidence and mortality of cervical cancer in 2020: a baseline analysis of the WHO Global Cervical Cancer Elimination Initiative. Lancet Global Health. 2023;11(2):e197–206.
3. Han B et al. Cancer incidence and mortality in China, 2022. J Natl Cancer Cent. 2024. https://doi.org/10.1016/j.jncc.2024.01.006
4. Okunade KS. Human papillomavirus and cervical cancer. J Obstet Gynaecol. 2020;40(5):602–8.
5. Chelimo C, Wouldes TA, Cameron LD, Elwood JM. Risk factors for and prevention of human papillomaviruses (HPV), genital warts and cervical cancer. J Infect. 2013;66(3):207–17.
6. Magalhães GM, et al. Update on human papilloma virus - part I: epidemiology, pathogenesis, and clinical spectrum. An Bras Dermatol. 2021;96(1):1–16.
7. Gao B, et al. The characteristics and risk factors of human papillomavirus infection: an outpatient population-based study in Changsha, Hunan. Sci Rep. 2021;11(1):15128.
8. Piyathilake CJ, et al. Cervical Microbiota Associated with higher Grade Cervical Intraepithelial Neoplasia in Women infected with high-risk human papillomaviruses. Cancer Prev Res. 2016;9(5):357–66.
9. Consolaro MEL, et al. Changes of vaginal microbiota during cervical carcinogenesis in women with human papillomavirus infection. PLoS ONE. 2020;15(9):e0238705.
10. Łaniewski P, et al. Linking cervicovaginal immune signatures, HPV and microbiota composition in cervical carcinogenesis in non-hispanic and hispanic women. Sci Rep. 2018;8(1):7593.
11. Dareng EO, et al. Prevalent high-risk HPV infection and vaginal microbiota in Nigerian women. Epidemiol Infect. 2015;144(1):123–37.
12. Mitra A, et al. Cervical intraepithelial neoplasia disease progression is associated with increased vaginal microbiome diversity. Sci Rep. 2015;5(1):16865.
13. Klein C et al. Relationship between the cervical microbiome, HIV Status, and precancerous lesions. mBio 2019; 10(1):e02785-18.
14. Clifford GM, et al. Worldwide distribution of human papillomavirus types in cytologically normal women in the International Agency for Research on Cancer HPV prevalence surveys: a pooled analysis. Lancet. 2005;366(9490):991–8.
15. Balachandra S, et al. Blood-based biomarkers of human papillomavirus–associated cancers: a systematic review and meta-analysis. Cancer. 2020;127(6):850–64.
16. Borgogna JC, et al. The vaginal metabolome and microbiota of cervical HPV-positive and HPV-negative women: a cross-sectional analysis. BJOG. 2020;127(2):182–92.
17. Chorna N, Romaguera J, Godoy-Vitorino F. Cervicovaginal Microbiome and urine metabolome paired analysis reveals niche partitioning of the Microbiota in patients with human papilloma virus infections. Metabolites 2020; 10(1).

18. Cheung LC, et al. 2019 ASCCP Risk-Based Management Consensus guidelines: methods for risk estimation, recommended management, and validation. J Lower Genit Tract Dis. 2020;24(2):90–101.

19. Yang Y, et al. Pueraria lobata starch regulates gut microbiota and alleviates high-fat high-cholesterol diet induced non-alcoholic fatty liver disease in mice. Food Res Int. 2022;157:111401.

20. Yang Y, et al. Starch from Pueraria lobata and the amylose fraction alleviates dextran sodium sulfate induced colitis in mice. Carbohydr Polym. 2023;302:120329.

21. Johnson JS, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat Commun. 2019;10(1):5029.

22. Imaizumi K, Tinwongger S, Kondo H, Hirono I. Analysis of microbiota in the stomach and midgut of two penaeid shrimps during probiotic feeding. Sci Rep. 2021;11(1):9936.

23. Shang T, et al. The combination of four main components in Xuebijing injection improved the preventive effects of cyclosporin A in acute graft-versus-host disease mice by protecting intestinal microenvironment. Biomed Pharmacother. 2022;148:112675.

24. Wu X, et al. An integrated microbiome and metabolomic analysis identifies immunoenhancing features of Ganoderma lucidum spores oil in mice. Pharmacol Res. 2020;158:104937.

25. Chen Y, et al. Human papillomavirus infection and cervical intraepithelial neoplasia progression are associated with increased vaginal microbiome diversity in a Chinese cohort. BMC Infect Dis. 2020;20(1):629.

26. Huang X, et al. Cervicovaginal microbiota composition correlates with the acquisition of high-risk human papillomavirus types. Int J Cancer. 2018;143(3):621–34.

27. Shannon B, et al. Association of HPV infection and clearance with cervicovaginal immunology and the vaginal microbiota. Mucosal Immunol. 2017;10(5):1310–9.

28. Abulizi G, et al. Risk factors for human papillomavirus infection prevalent among uyghur women from Xinjiang, China. Oncotarget. 2017;8(58):97955–64.

29. Borgdorff H, et al. Lactobacillus-dominated cervicovaginal microbiota associated with reduced HIV/STI prevalence and genital HIV viral load in African women. ISME J. 2014;8(9):1781–93.

30. Feng R-M, et al. Risk of high-risk human papillomavirus infection and cervical precancerous lesions with past or current trichomonas infection: a pooled analysis of 25,054 women in rural China. J Clin Virol. 2018;99–100:84–90.

31. Younes JA, et al. Women and their microbes: the unexpected friendship. Trends Microbiol. 2018;26(1):16–32.

32. Kovachev S. Defence factors of vaginal lactobacilli. Crit Rev Microbiol. 2017;44(1):31–9.

33. Samar Megbal A, Sherif E, Ahmed B, Rashad Rizk A-H. The composition and stability of the vaginal microbiome of healthy women. J Pak Med Assoc. 2021;71(8):1–20.

34. Brotman RM, et al. Association between the vaginal microbiota, menopause status, and signs of vulvovaginal atrophy. Menopause. 2018;25(11):1321–30.

35. Gillet E, et al. Bacterial vaginosis is associated with uterine cervical human papillomavirus infection: a meta-analysis. BMC Infect Dis. 2011;11(1):10.

36. Di Paola M, et al. Characterization of cervico-vaginal microbiota in women developing persistent high-risk human papillomavirus infection. Sci Rep. 2017;7(1):10200.

37. Frąszczak K, Barczyński B, Kondracka A. Does Lactobacillus Exert a Protective Effect on the Development of Cervical and Endometrial Cancer in Women? Cancers 2022; 14(19).

38. Fan Q, et al. Lactobacillus spp. create a protective micro-ecological environment through regulating the core fucosylation of vaginal epithelial cells against cervical cancer. Cell Death Dis. 2021;12(12):1094.

39. Aroutcheva A, et al. Defense factors of vaginal lactobacilli. Am J Obstet Gynecol. 2001;185(2):375–9.

40. Nishiumi F, et al. Blockade of endoplasmic reticulum stress-induced cell death by Ureaplasma parvum vacuolating factor. Cell Microbiol. 2021;23(12):e13392.

41. Eisenberg T et al. Sneathia vaginalis sp. nov. (Fusobacteriales, Leptotrichiaceae) as a replacement of the species 'Sneathia amnii' Harwich 2012 and 'Leptotrichia amnionii' Shukla 2002, and emended description of Sneathia Collins 2001. International Journal of Systematic and Evolutionary Microbiology 2021; 71(3):004663.

42. Harwich MD, et al. Genomic sequence analysis and characterization of Sneathia amnii sp. nov. BMC Genomics. 2012;13(S8):S4.

43. Adebamowo SN et al. Mycoplasma hominis and Mycoplasma genitalium in the Vaginal Microbiota and Persistent High-Risk Human Papillomavirus Infection. Frontiers in Public Health 2017; 5:140.

44. Salliss ME, Maarsingh JD, Garza C, Łaniewski P, Herbst-Kralovetz MM. Veillonellaceae family members uniquely alter the cervical metabolic microenvironment in a human three-dimensional epithelial model. Npj Biofilms Microbiomes. 2021;7(1):57.

45. Carlstein C, Marie Søes L, Jørgen Christensen J. Aerococcus christensenii as part of severe Polymicrobial Chorioamnionitis in a pregnant woman. Open Microbiol J. 2016;10(1):27–31.

46. Deshpande R, Mansara P, Kaul-Ghanekar R. Alpha-linolenic acid regulates Cox2/VEGF/MAP kinase pathway and decreases the expression of HPV oncoproteins E6/E7 through restoration of p53 and rb expression in human cervical cancer cell lines. Tumor Biology. 2015;37(3):3295–305.

47. Klein V, et al. Low alpha-linolenic acid content of adipose breast tissue is associated with an increased risk of breast cancer. Eur J Cancer. 2000;36(3):335–40.

48. Bougnoux P, et al. α-Linolenic acid content of adipose breast tissue: a host determinant of the risk of early metastasis in breast cancer. Br J Cancer. 1994;70(2):330–4.

49. Moon H-S, Batirel S, Mantzoros CS. Alpha linolenic acid and oleic acid additively down-regulate malignant potential and positively cross-regulate AMPK/S6 axis in OE19 and OE33 esophageal cancer cells. Metabolism. 2014;63(11):1447–54.

50. Reimann B, et al. In utero exposure to parabens and early childhood BMI z-scores – associations between placental ethyl paraben, longitudinal BMI trajectories and cord blood metabolic biomarkers. Environ Int. 2021;157:106845.

51. Zhang G, et al. UPLC-Q-TOF/MS-based plasma metabolome to identify biomarkers and time of injury in traumatic brain injured rats. NeuroReport. 2021;32(6):415–22.

52. Edenharder R, Slemrova J. The significance of the bacterial steroid degradation for the etiology of large bowel cancer. IV. Deconjugation of glycocholic acid, oxidation, and reduction of cholic acid by saccharolytic Bacteroides species. Zentralbl Bakteriol Orig B. 1976;162:350–73.

53. Viswanathan S, et al. Involvement of CYP 2C9 in mediating the Proinflammatory effects of Linoleic Acid in Vascular endothelial cells. J Am Coll Nutr. 2003;22(6):502–10.

54. Tolstikov V, Moser AJ, Sarangarajan R, Narain NR, Kiebish MA. Current status of Metabolomic Biomarker Discovery: impact of Study Design and demographic characteristics. Metabolites. 2020;10(6):224.

Publisher's Note