The association of uterine cervical microbiota with an increased risk for cervical intraepithelial neoplasia in Korea

H. Y. Oh1, B.-S. Kim2,3, S.-S. Seo4, J.-S. Kong1, J.-K. Lee5, S.-Y. Park4, K.-M. Hong6, H.-K. Kim6 and M. K. Kim1

1) Division of Cancer Epidemiology and Management, National Cancer Centre, Ilsandong-gu, Goyang, 2) Chunlab Inc., Seoul National University, Seoul, 3) Department of Life Science, Hallym University, Chuncheon, 4) Centre for Uterine Cancer, National Cancer Centre, Ilsandong-gu, Goyang, 5) Department of Obstetrics and Gynaecology, Korea University College of Medicine, Seoul and 6) Division of Cancer Biology, National Cancer Centre, Goyang-si

Abstract

Recent studies have suggested potential roles of the microbiome in cervicovaginal diseases. However, there has been no report on the cervical microbiome in cervical intraepithelial neoplasia (CIN). We aimed to identify the cervical microbiota of Korean women and assess the association between the cervical microbiota and CIN, and to determine the combined effect of the microbiota and human papillomavirus (HPV) on the risk of CIN. The cervical microbiota of 70 women with CIN and 50 control women was analysed using pyrosequencing based on the 16S rRNA gene. The associations between specific microbial patterns or abundance of specific microbiota and CIN risk were assessed using multivariate logistic regression, and the relative excess risk due to interaction (RERI) and the synergy index (S) were calculated. The phyla Firmicutes, Actinobacteria, Bacteroidetes, Proteobacteria, Tenericutes, Fusobacteria and TM7 were predominant in the microbiota and four distinct community types were observed in all women. A high score of the pattern characterized by predominance of Atopobium vaginae, Gardnerella vaginalis and Lactobacillus iners with a minority of Lactobacillus crispatus had a higher CIN risk (OR 5.80, 95% CI 1.73–19.4) and abundance of A. vaginae had a higher CIN risk (OR 6.63, 95% CI 1.61–27.2). The synergistic effect of a high score of this microbial pattern and oncogenic HPV was observed (OR 34.1, 95% CI 4.95–284.5; RERI/S, 15.9/1.93). A predominance of A. vaginae, G. vaginalis and L. iners with a concomitant paucity of L. crispatus in the cervical microbiota was associated with CIN risk, suggesting that bacterial dysbiosis and its combination with oncogenic HPV may be a risk factor for cervical neoplasia.

Clinical Microbiology and Infection © 2015 European Society of Clinical Microbiology and Infectious Diseases. Published by Elsevier Ltd. All rights reserved.

Keywords: Atopobium vaginae, cervical intraepithelial neoplasia, human papillomavirus, Lactobacillus crispatus, Lactobacillus iners, microbiome, pyrosequencing

Original Submission: 22 September 2014; Revised Submission: 4 January 2015; Accepted: 26 February 2015

Editor: D. Raoult

Article published online: 6 March 2015

Introduction

The importance of the microbiota in health maintenance or disease susceptibility is well known; the gut microbiota is known to control the absorption of nutrients [1]. Imbalance in the gut is associated with Crohn’s disease, type 2 diabetes and chronic allergies [2], for which dysbiosis, not a single pathogen change but shifts in the relative abundance of microbes, has long been suggested. Furthermore, the association of microbial dysbiosis with several cancer types has been suggested in areas surrounded by mucosal membranes where
bacteria live densely [3]. Although only modest effects of dysbiosis on cancer have been presented so far, the long duration of dysbiosis and possible combined effects with other risk factors suggest greater clinical implications [4]. Although human papillomavirus (HPV) has been the major risk for cervical precancerous lesions or cancer [5]; recently, the potential role of the cervicovaginal microbiome in cervical cancer through the elevation of pH has been reported [6]. In addition, the role of the cervicovaginal microbiome in HPV infection has also been reported [7], suggesting a possible role in cervical cancer through potentiation of HPV infection. Although there have been reports on the microbiome in premenopausal or postmenopausal women [8], there are still no studies on the microbiome related to cervical precancerous lesions or cancers being presented. The purpose of this study was to assess the association between cervical intraepithelial neoplasia (CIN) and cervical microbiota identified using pyrosequencing.

**Materials and methods**

**Study participants and design**

The Korean HPV cohort study, which was established to identify epidemic, genetic, viral and ecological factors associated with the development of cervical neoplasia in Korean women, recruited 1096 women, aged 18–65 years, from March 2006 to the present. The study participants were randomly selected from the gynaecological oncology clinics of six tertiary medical centres in Korea. Detailed information regarding the inclusion/exclusion criteria and the design of the baseline measurements for the HPV cohort were presented in a previous study [10]. All study participants provided written informed consent in accordance with good clinical practices. Cervical swabs were collected for a Papacinoacolau smear test. Immediately after sampling, each cervix brush (Rovers Medical Devices, Oss, the Netherlands) was rinsed in a vial of PreservCyt solution (Cytyc Co., Marlborough, MA, USA), and the vial was placed in a Thin Prep Processor (Cytyc Co., Gaithersburg, MD, USA). Half of the second swab was collected for high-risk (HR)-HPV DNA detection and microbiota analysis using a Cervical Sampler Brush (Digene Co. Gaithersburg, MD, USA). Half of the second swab was used immediately for HR-HPV DNA detection and the other half was stored at −80°C for subsequent DNA extraction, for 2 months up to a maximum of 5 years. A total of 125 women were randomly selected from the 1096 enrolled women, and baseline samples obtained from 120 women (50 controls, 70 with CIN) were analysed; for five of the samples, metagenomic DNA extraction failed. This study was approved by the Institutional Review Board of the Korean National Cancer Centre (NCCNCS-06-062) and by the ethics committees of the Korean National Cancer Centre and Korea University Guro Hospital.

**Cytological screening and HR-HPV DNA detection**

The cervical cytological findings were classified according to the Bethesda system [11]. HPV DNA detection was performed using the commercially available Digene Hybrid capture II DNA Test (Qiagen, Gaithersburg, MD, USA). The results of a chemiluminescent HPV DNA test were measured in relative light units (RLU) with a probe designed to detect 13 types of HR-HPV. The test results were read as positive at concentrations of 1 pg/mL or greater than the RLU/cut-off ratio (RLU of specimen/mean RLU of two positive controls).

**DNA extraction and pyrosequencing**

Metagenomic DNA samples were extracted using the Fast DNA SPIN extraction kits (MP Biomedicals, Santa Ana, CA, USA). Target fragments of the 16S rRNA gene corresponding to the V1–V3 regions were amplified using bar-coded primers. Amplifications were performed in a final volume of 50 μL containing 10 × Taq buffer, a dNTP mixture (Takara, Shiga, Japan), 10 μM of the bar-coded fusion primers, and 2 U of Taq polymerase (ExTaq, Takara). The amplification conditions and pyrosequencing procedure have been previously described [12]. The beads recovered following emulsion PCR were deposited on a 454 Pico Titer Plate, and sequencing was performed using a Roche/454 GS Junior system (Roche, Branford, CT, USA).

**Pyrosequencing data analysis**

The raw sequences were sorted using their unique barcodes, and low-quality reads (with an average quality score <25 or a read length <300 bp) were removed [12]. The primer sequences were trimmed using a pairwise sequence alignment, and sequences were clustered to correct for sequencing errors. Representative sequences in each cluster were identified using EzTaxon-e, a public database that contains sequences of the 16S rRNA genes of species type strains, and the taxonomic positions of relevant uncultured sequences were identified [13] using the highest pairwise similarity among the BLASTN results. Chimeric sequences were detected and removed by the UCHIME algorithm, which detects chimeric sequences by alignment of a query sequence with two parent sequences in a reference database [14], and the diversity indices were calculated using the Mothur program. The heat-map analysis was performed using the Multi-EXPERIMENT VIEWER. Pyrosequencing reads are available in the EMBL SRA database (http://www.ebi.ac.uk/ena/data/view/PRJEB5760).
Quantitative real-time PCR assays
Genomic DNA was extracted with 25 KU/mL mutanolysin (Sigma-Aldrich, St Louis, MO, USA) using a QIAamp DNA mini kit (Qiagen, Valencia, CA, USA), as described previously [15]. The DNA targets, primers, probe sequences, oligonucleotides used for PCR, and the amplification conditions are shown in the Supporting information (Table S1) [16].

Statistical analysis
Factor analysis was performed on the 18 most predominant species using STATA 12.0 (Stata Co., College Station, TX, USA). Among the 18 microbial patterns identified, ten had an eigenvalue >1. The pattern matrix (i.e. the correlation coefficients between the species and the patterns) is provided in the Supporting information (Table S2). The Wilcoxon rank-sum and Spearman’s rank correlation tests were performed to measure differences between two groups and the dependence between two variables, respectively. Logistic regression analysis was performed after adjustment for age, marital status, menopausal status, oral contraceptive use, smoking habit and HR-HPV infection. Risk estimates are presented as OR with 95% CI. To determine biological interactions, ORs for additive scale and multiplicative terms were estimated, and the relative excess risk due to interaction (RERI) and synergy index (S) were calculated as described previously [17]. Values for RERI >0 and S >1 indicate the presence of a synergistic effect.

Results

General characteristics of study participants
Women with CIN 1 (n = 55) and CIN 2 or 3 (n = 15) were accrued, along with control women with normal cytology.
(n = 25) or atypical squamous cells of undetermined significance (ASCUS; n = 25) (see Supporting information, Table S2). No significant difference was observed between control women and women with CIN except HR-HPV infection (p = 0.002).

**Cervical microbiota among participants**

In total, 1,118,398 high-quality reads were analysed and an average of 74.1 operational taxonomic units (OTUs) per sample was observed. Seven phyla—Firmicutes, Actinobacteria, Bacteroidetes, Proteobacteria, Tenericutes, Fusobacteria, and candidate division TM7—were dominant (see Supporting information, Fig. S1). Although Firmicutes was the predominating phylum in most women (n = 75), Bacteroidetes, Actinobacteria, Tenericutes and Proteobacteria were also prevalent (n = 45). The composition of the dominant phyla was not clearly distributed by HPV infection status or histological grades. However, the abundances of Bacteroidetes, Actinobacteria, Tenericutes and Proteobacteria were higher in HPV-positive women (n = 34) than in HPV-negative women (14). Actinobacteria was observed more predominantly in women with CIN (n = 5) than in controls (6).

The OTU number was higher in women with CIN than controls in postmenopausal women, but lower in women with CIN than controls in premenopausal women (see Supporting information, Fig. S2a). The OTU number was higher in HPV-negative than HPV-positive women irrespective of menopausal status (see Supporting information, Fig. S2b). The OTU number in control women and women with CIN was similar in HPV-negative women, but it was higher in controls than in women with CIN in HPV-positive women (see Supporting information, Fig. S2c).

Cervical microbiota was distributed into four clusters using heat-map analysis at the species level (Fig. 1). *Lactobacillus iners* was predominant in cluster I; *Atopobium vaginae*, *Prevotella bivia*, *Lactobacillus fornicalis*, *Pseudomonas poae* and *Gardnerella vaginalis* were dominant members in cluster II; *L. iners* and *Lactobacillus crispatus* were dominant in cluster III; and *L. iners* and *L. crispatus* were dominant in cluster IV. The clusters were not clearly separated by HPV infection or histology grade, most HPV-negative women carried *L. iners* and *L. crispatus* as the dominant members. There were more HPV-positive women in clusters II and IV, and *A. vaginae*, *P. bivia*, *L. fornicalis*, *P. poae* and *G. vaginalis* were higher in cluster II than in cluster IV.

**Risky microbial pattern associated with CIN risk**

The individual scores for the ten microbial patterns each obtained from factor analysis were used for the following step. Only the third pattern showed a significant difference in its score between controls and women with CIN (p = 0.007) (Table 1 and see Supporting information, Table S3). The OR of the tertiles and p < 0.05 was regarded as significant.
G. vaginalis and a concomitant paucity of L. crispatus (see Supporting information, Table S4). The women with a high tertile of this pattern score had a high proportion of A. vaginae (mean 28.8%) and G. vaginalis (5.6%), and a low proportion of L. crispatus (0.95%) (Fig. 2a). Lactobacillus iners predominated in women with the medium to high scores (21.2%) (Fig. 2b).

Association of A. vaginae, G. vaginalis, L. iners and L. crispatus with CIN risk

There was a significant median difference in relative abundance of A. vaginae between controls and women with CIN (p 0.003) (Table 1). The OR for CIN risk of a high tertile of A. vaginae relative abundance was 6.63 (95% CI 1.61–27.2) compared with the low tertile. The abundances of L. iners and L. crispatus were not different between controls and women with CIN, and high tertiles of those species were not associated with CIN. However, some women with CIN had a high proportion of L. iners (Fig. 1) and some had complete absence of or low L. crispatus.

Synergistic effect of the risky microbial pattern and HR-HPV on CIN risk

The OR for CIN risk of HR-HPV-positive women with a high score of the risky microbial pattern was 34.1 (95% CI 4.95–234.5), compared with HR-HPV-negative women with a low score. The RERI and S (15.9 and 1.93), and p value for the multiplicative term (p 0.018) also showed a synergistic effect between the two (Table 2). The OR of HR-HPV-positive women with A. vaginae in the cervix was 29.9 (95% CI 5.49–163.0), compared with HR-HPV-negative women without A. vaginae (RERI and S, 15.4 and 2.14; p 0.002). The OR of HR-HPV-positive women with L. iners in the cervix was 10.8 (95% CI 1.65–70.9) (RERI and S, 7.39 and 4.06; p < 0.001).

Quantitative real-time PCR for A. vaginae and L. crispatus

To compare the results obtained from the quantitative real-time PCR assay and the pyrosequencing for A. vaginae and
Synergistic joint effect of the risky microbial pattern with high-risk human papillomavirus (HR-HPV) infection on the increase of cervical intraepithelial neoplasia risk

In most women, L. crispatus and L. iners were the dominant cervical microbiota. These two species were also found to be predominant in previous studies of the vaginal microbiota of Asian women [8,18]. In contrast to previous reports, A. vaginae, G. vaginalis, P. bivia and Pneumococcus poae were the dominant species in several women in this study. In previous reports, a higher proportion of Proteobacteria, a lower proportion of Firmicutes, and a greater diversity of bacteria were observed in postmenopausal women than in premenopausal women [8]. However, in the present study, various microbial compositions and a higher proportion of Firmicutes were observed in postmenopausal women. In addition, higher numbers of OTU were detected in premenopausal women than in postmenopausal women. This difference may be attributed to differences in the target region of 16S rRNA genes analysed or in sequencing methods [19]. Although these differences could produce disparities among the results of studies of microbiota, the predominant members of the cervical microbiota detected here were consistent with those detected in previous studies that used different target regions [19]. The similarity of the results is a result of the relatively low diversity of the cervical microbiota, compared with microbiota from other parts of the human body. Therefore, various other factors, including sample collection techniques, temporal shifts in the microbiota during the menstrual cycle, hygiene practices, glycogen levels and host genetic factors could be associated with individual variation; these variations may underlie the discrepancy observed between the results of this study and previous studies [8,18–21].

**TABLE 2. Synergistic joint effect of the risky microbial pattern with high-risk human papillomavirus (HR-HPV) infection on the increase of cervical intraepithelial neoplasia risk**

<table>
<thead>
<tr>
<th>Joints</th>
<th>Symbol</th>
<th>N w/wo</th>
<th>OR (95% CI)</th>
<th>RERI/S</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Risky microbial pattern (a) and HR-HPV (b)</strong></td>
<td>0, 0</td>
<td>5/19</td>
<td>1 (ref)</td>
<td>15.9/1.93</td>
</tr>
<tr>
<td>p for trend</td>
<td>0.120</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p for interaction (a × b)</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Atoxopibium vaginae (a) and HR-HPV (b)</strong></td>
<td>0, 0</td>
<td>4/15</td>
<td>1 (ref)</td>
<td>15.4/2.14</td>
</tr>
<tr>
<td>p for trend</td>
<td>0.031</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p for interaction (a × b)</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gardnerella vaginalis (a) and HR-HPV (b)</strong></td>
<td>0, 0</td>
<td>7/18</td>
<td>1 (ref)</td>
<td>1.23/1.19</td>
</tr>
<tr>
<td>p for trend</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p for interaction (a × b)</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Lactobacillus iners (a) and HR-HPV (b)</strong></td>
<td>0, 0</td>
<td>2/6</td>
<td>1 (ref)</td>
<td>7.39/4.06</td>
</tr>
<tr>
<td>p for trend</td>
<td>0.280</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p for interaction (a × b)</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Lactobacillus crispatus (a) and HR-HPV (b)</strong></td>
<td>1, 0</td>
<td>7/14</td>
<td>1 (ref)</td>
<td>−2.28/0.66</td>
</tr>
<tr>
<td>p for trend</td>
<td>0.250</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*aThe symbol ‘0’ corresponds to the low score of the risky microbial pattern (low to medium tertile), the absence of indicated microbial species (the relative abundance = 0) and non-HPV infection. The symbol ‘1’ stands for the high risky microbial pattern score (high tertile), the presence of that (the relative abundance > 0) and HPV infection.

*bThe number of women with/without cervical intraepithelial neoplasia (CIN) lesions was presented.

*cOdds ratio (OR) was estimated after adjustment for age, marital status, menopausal status, oral contraceptive use and smoking history as categorical variables.

*dThe biological interaction of the joint based on the additive model was investigated using relative excess risk due to interaction (RERI) and synergy index (S). RERI >0 and S > 1 indicate a synergistic effect.

Discussion

This study showed that a cervical microbial composition characterized by a predominance of A. vaginae, G. vaginalis and L. iners with a concomitant paucity of L. crispatus is associated with a higher risk of CIN. The predominance of A. vaginae was a major contributor to this risk. Furthermore, a synergistic interaction between this microbial pattern and HR-HPV infection that robustly increased the risk was observed.

**Clinical Microbiology and Infection © 2015 European Society of Clinical Microbiology and Infectious Diseases. Published by Elsevier Ltd. All rights reserved, CMI, 21, 674.e1–674.e9**
Clusters I and III contained a greater abundance of *L. iners* and *L. crispatus*. Clusters II and IV contained various bacteria as the predominant members, but differed in composition. These clusters were similar to previously reported community state types [18]. Cluster II was similar to state type IV_B, and cluster IV was similar to type IV_A. Diverse microbial associations containing high proportions of *Atopobium*, *Prevotella*, *Sneathia* and *Gardnerella* have been reported to be related to BV [22]. Although the diverse community detected in clusters II and IV could be related to BV, diverse bacterial communities do not always have persistently high Nugent scores [21]. A high Nugent score is one of the methods available for diagnosis of BV, but it is not routinely used by physicians because of its inaccuracy. Molecular quantification of *A. vaginae* and *G. vaginalis* shows a higher reproducibility for diagnosis of BV than Nugent scores [16].

Although confirmation using other diagnostic tests is required, BV may be involved in the microbial composition associated with a high risk of CIN identified in the study. To date, several studies have explored the role of BV in cervical neoplasia; they have reported conflicting findings. In one prospective study, the CIN rate and the quantity of nitrosamines produced in women with BV did not differ from those in women with normal flora [23]. Another longitudinal study reported no influence of abnormal flora on the presence of atypical endocervical cells [24]. In contrast, Nam et al. showed a significant correlation between BV and CIN, regardless of severity [25]. Furthermore, a recently published systemic review and meta-analysis showed a positive correlation between BV and CIN [26]. However, these ambivalent results may be attributed to the different tools used in the diagnosis of BV. Of the above-mentioned studies, two used a BV diagnosis with three other measurements, including pH, clue cells after Gram stain, and a positive amine or whiff test [23,25], and the other used the Nugent scale [24]. Although the correlation between BV and CIN remains to be determined, our results are in support of such a correlation.

In this study, the women with the highest risk of CIN showed low levels of *L. iners* and *L. crispatus*. Hydrogen peroxide-producing *Lactobacilli* are present in 96% of women with a normal vaginal bacterial community [27]. Species in the genus *Lactobacillus* maintain a low pH by producing lactic acid. *Lactobacillus crispatus* was more strongly associated with a low vaginal pH than *L. iners*, suggesting that these two species differ in ecological function [9]. Whereas a high abundance of *L. crispatus* was detected in the low-risk group in this study, *L. iners* was abundant in the medium-risk group, showing that this population indicates a susceptibility to CIN.

*Atopobium vaginae* also stimulates an innate immune response from vaginal epithelial cells, leading to the secretion of interleukin, and the production of defending protein β-defensin (BD) [30]. Increased protein and mRNA expression of human BD2 and human BD3 was also observed in the genital condylomata acuminata of the vulvovaginal tract, induced by HPV.

![FIG. 3. Quantitative real-time PCR assay of *Atopobium vaginae* and *Lactobacillus crispatus*. Log of the ratio of *A. vaginae* and *L. crispatus* provided from pyrosequencing (the relative abundances) (a) and real-time PCR (equivalent/μL) (b). A Spearman’s rank correlation test was performed for measuring the statistical dependence between two variables; rank correlation coefficient = 0.868, p < 0.001. The p value was from the Wilcoxon rank-sum test for controls and women with cervical intraepithelial neoplasia (CIN).](image-url)
Colonization of the vaginal epithelia by A. vaginae and G. vaginalis induced production of interleukin-8 and regulated on activation, normal T-cell expressed and secreted (RANTES), and amplified the pro-inflammatory response to both membrane lipophosphoglycan/ceramide–phosphoinositol–glycan core and Trichomonas vaginalis [32]. Levels of RANTES and macrophage inflammatory protein-1β were significantly higher in the cervical mucosa of HIV/HPV-co-infected individuals than in HPV mono-infected individuals [33]. In a recent survey of the interactions between HPV types in healthy individuals, more than 50% of healthy women showed coinfection with two to three HPV types in vaginal tissue [34]. Evidence of complex interactions between viruses and immune regulators, bacterial flora and immune regulators, and co-infection with multiple viruses in the genital tract is accumulating [30–34]. Our findings demonstrated an interaction between the bacterial flora and HR-HPV-type infections, and a possible association between this interaction and clinical cervical neoplasia. The synergistic effect that increases the risk of CIN noted between a predominance of BV-related bacteria and a concomitant HPV infection could be caused by a robust response of the same immune regulators.

This study demonstrated that bacterial dysbiosis, characterized by a predominance of A. vaginae, G. vaginalis and L. iners and a concomitant paucity of L. crispatus, in its combination with oncogenic HPV may be a risk factor for cervical neoplasia. Although further studies using a large number of women with CIN 2 or CIN 3 and considering other latent sexually transmitted organisms with no symptoms are required, we suggest that a monitoring programme for the relative distribution of these species in the cervicovaginal system could be helpful for the effective prevention of cervical precancerous lesions.

Transparency declaration

The authors have no conflict of interest.

Acknowledgements

This research was supported by grants from the National Cancer Centre in Korea (1110320, 1310360).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.cmi.2015.02.026

References


