



# Relationship between the Cervical Microbiome, HIV Status, and Precancerous Lesions

Cameron Klein,<sup>a</sup> Daniela Gonzalez,<sup>a</sup> Kandali Samwel,<sup>a</sup> Crispin Kahesa,<sup>b</sup> Julius Mwaiselage,<sup>b</sup> Niros Aluthge,<sup>c</sup> Samodha Fernando,<sup>c</sup> John T. West,<sup>a</sup> Charles Wood,<sup>a</sup> Peter C. Angeletti<sup>a</sup>

<sup>a</sup>Nebraska Center for Virology, School of Biological Sciences, University of Nebraska-Lincoln, Lincoln, Nebraska, USA

<sup>b</sup>Ocean Road Cancer Institute, Dar es Salaam, Tanzania

<sup>c</sup>Animal Sciences, University of Nebraska-Lincoln, Lincoln, Nebraska, USA

**ABSTRACT** Nearly all cervical cancers are causally associated with human papillomavirus (HPV). The burden of HPV-associated dysplasias in sub-Saharan Africa is influenced by HIV. To investigate the role of the bacterial microbiome in cervical dysplasia, cytobrush samples were collected directly from cervical lesions of 144 Tanzanian women. The V4 hypervariable region of the 16S rRNA gene was amplified and deep sequenced. Alpha diversity metrics (Chao1, PD whole tree, and operational taxonomic unit [OTU] estimates) displayed significantly higher bacterial richness in HIV-positive patients ( $P = 0.01$ ) than in HIV-negative patients. In HIV-positive patients, there was higher bacterial richness in patients with high-grade squamous intraepithelial lesions (HSIL) ( $P = 0.13$ ) than those without lesions. The most abundant OTUs associated with high-grade squamous intraepithelial lesions were *Mycoplasma* spp., *Pseudomonadales*, and *Staphylococcus*. We suggest that a chronic mycoplasma infection of the cervix may contribute to HPV-dependent dysplasia by sustained inflammatory signals.

**IMPORTANCE** HPV is known to be the causal agent in the majority of cervical cancers. However, the role of the cervical bacterial microbiome in cervical cancer is not clear. To investigate that possibility, we collected cervical cytobrush samples from 144 Tanzanian women and performed deep sequencing of bacterial 16S rRNA genes. We found that HIV-positive patients had greater bacterial richness ( $P = 0.01$ ) than HIV-negative patients. We also observed that women with high-grade squamous intraepithelial lesions (HSIL) had greater cervical bacterial diversity than women with cytologically normal cervixes. Data from our precise sampling of cervical lesions leads us to propose that *Mycoplasma* contributes to a cervical microbiome status that promotes HPV-related cervical lesions. These results suggest a greater influence of the bacterial microbiota on the outcome of HPV infection than previously thought.

**KEYWORDS** 16S RNA, cervical cancer, deep sequencing, human papillomavirus, human immunodeficiency virus, microbiome

Human papillomavirus (HPV) is the causative agent responsible for 99% of cervical cancers (1). HPV contributes to about 4.8% of all cancers (1). The disease burden of HPV is most dramatic in developing regions of the world, with HPV contributing to 14.2% of cancers in sub-Saharan Africa (1). Cervical cancer disproportionately affects sub-Saharan Africa, where 9% of the world's female population over 15 years old accounts for 14% of the world's incidence of cervical cancer and 18% of cervical cancer-related deaths (2). The current study uses cervical swab samples obtained from Tanzania, which has among the highest cervical cancer mortality rates by country.

Sub-Saharan Africa also has among the highest HIV rates in the world. The associ-

**Citation** Klein C, Gonzalez D, Samwel K, Kahesa C, Mwaiselage J, Aluthge N, Fernando S, West JT, Wood C, Angeletti PC. 2019. Relationship between the cervical microbiome, HIV status, and precancerous lesions. mBio 10:e02785-18. <https://doi.org/10.1128/mBio.02785-18>.

**Invited Editor** Susan Hafenstein, The Pennsylvania State University College of Medicine

**Editor** Scott J. Hultgren, Washington University School of Medicine

**Copyright** © 2019 Klein et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Peter C. Angeletti, [Pangeletti2@unl.edu](mailto:Pangeletti2@unl.edu).

**Received** 11 December 2018

**Accepted** 8 January 2019

**Published** 19 February 2019

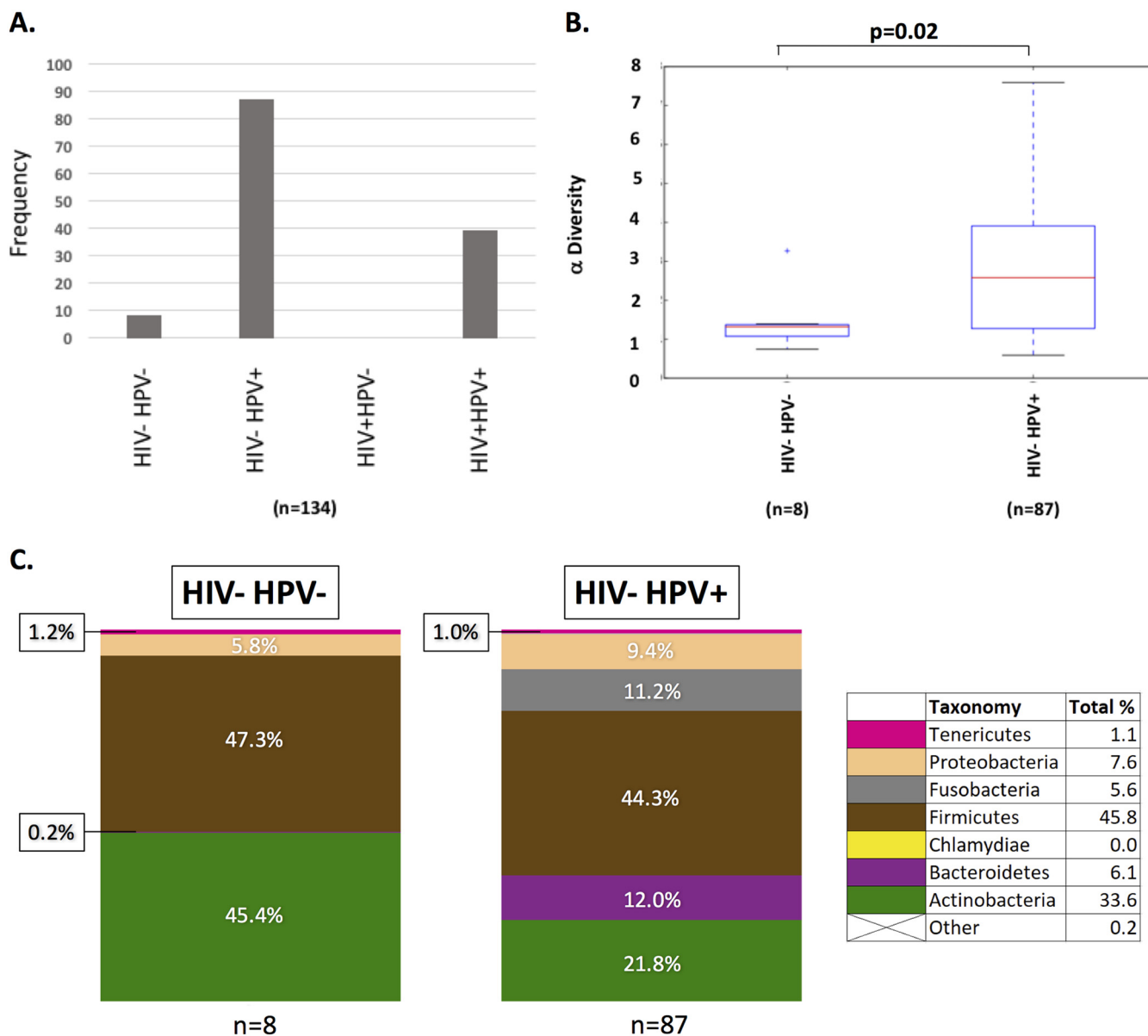
ation between HIV and cervical cancer has been better studied than any other factor associated with HPV-related cancers. HIV infection has been strongly linked to increased risk of infection with HPV and the severity of HPV pathogenesis (3–5). High-risk HPV genotypes are more prevalent in HIV-positive (HIV+) women, suggesting that HIV infection provides an environment where these high-risk HPVs can better establish infection and replicate (6). A likely factor in this is a decrease in T-cell surveillance, which results in an increase in HPV replication with decreasing CD4<sup>+</sup> cell count, and other changes in the cervical immune microenvironment as HIV infection progresses. Multiple studies have shown an increase in HPV detection in cervical intraepithelial neoplasms in individuals with less than 200 CD4<sup>+</sup> cells per  $\mu$ l of serum (7–10). Thus, the cervical immune microenvironment may be a cofactor in suppression of cervical cancer.

Changes in the cervicovaginal bacterial microbiome have been suggested to contribute to the development of precancerous cervical lesions (11–17). Chronic inflammation of the cervix (cervicitis), which is a result of cervicovaginal pathogens, leads to conditions like pelvic inflammatory disease (PID) and bacterial vaginosis (BV), both of which are associated with persistent HPV infection and cervical cancer (18, 19). Both PID and BV are more prevalent in sub-Saharan Africa and in HIV-positive populations (20–22). Comparative genomic analyses in women infected with HIV have shown that a shift in microbial diversity as a result of BV is detectable; whether this shift directly affects formation of precancerous cervical lesions is not clear (23). Given that cervical cancer rates are expected to rise in sub-Saharan Africa as the HIV-positive population receives life-extending antiretroviral therapy (ART), it is even more important to understand the risk factors associated with the cervical microbiome. There are previous studies that have analyzed how cervical microbiota differ at different stages of cervical cytology or as a function of HIV status (24–28). The current study defines bacterial communities associated with cervical lesions and with HIV, which represents a significant advance. Cervical cytology is graded by pap smear screening for nuclear abnormalities according to the Bethesda guidelines.

In this study, we utilized 16S rRNA gene deep sequencing on a set of 144 cervical swab samples from a cohort of Tanzanian women to gain an understanding of the differences in the cervical bacterial community composition as a function of cervical cytology grade and HIV status. The data presented here identify bacterial taxonomies associated with high-grade cervical lesions. In these studies, cervical lesions were sampled directly by cytobrush, instead of cervicovaginal lavage sampling. The rationale behind this approach was that the sites of the lesions are where tumors form, thus bacteria associated with lesion sites are more likely to be relevant to the process of disease progression than those associated with other regions.

## RESULTS

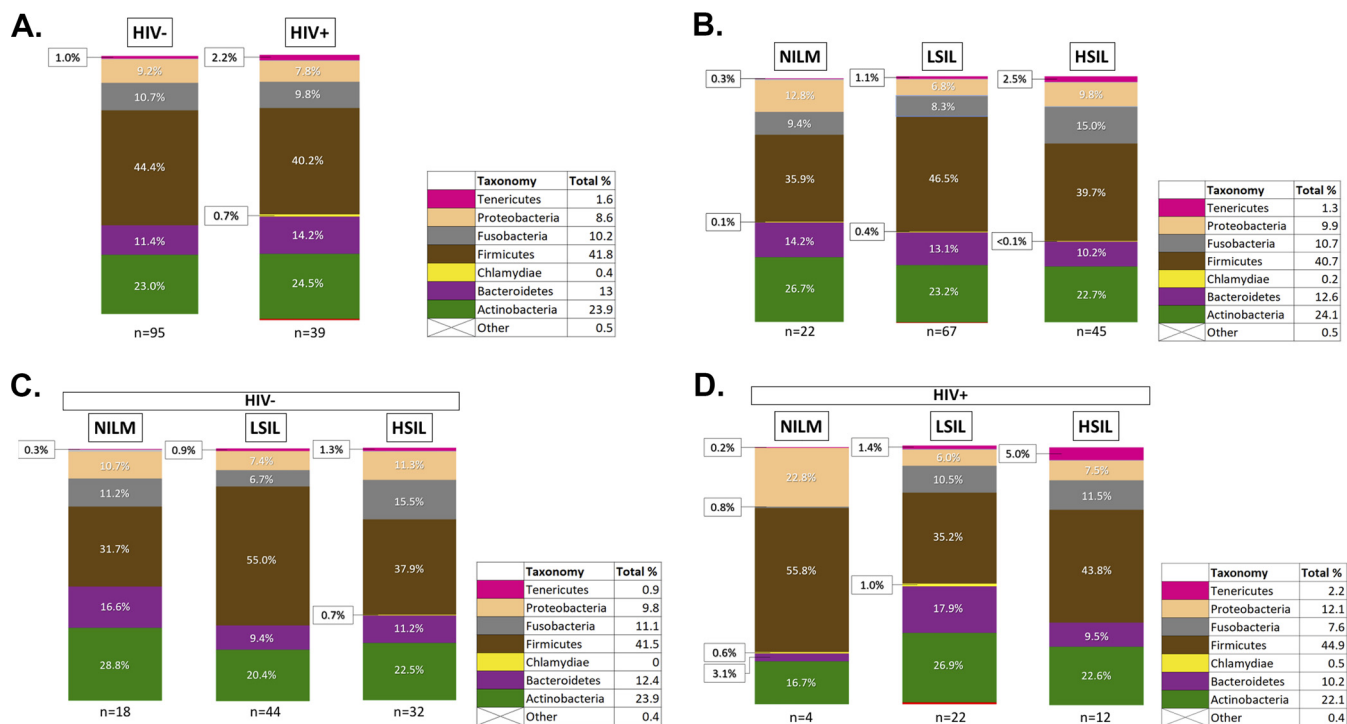
**Demographics.** Of the 144 patient samples, 41 were HIV positive (HIV+) and 103 were HIV negative (HIV–), with an average patient age of 37 years old. Of these 144 samples, 134 had HPV tests and deep sequencing reads of >1,000. The frequencies of HPV+ and HPV– samples with respect to HIV status are plotted in Fig. 1A. There were 8 HIV– HPV– samples and 87 HIV– HPV+ samples, but there were no HIV+ HPV– samples and 39 HIV+ HPV+ samples. Among HIV– samples, HPV had a statistically significant effect ( $P = 0.02$ ) on the cervical microbiome (Fig. 1B and C). Those microbes which were enriched in HPV+ samples were *Bacteroidetes* and fusobacteria. Also, there was a decrease in *Actinobacteria*. Cervical cytology was determined to be negative for intraepithelial lesion or malignancy (NILM) in 23 samples, low-grade squamous intraepithelial lesions (LSIL) in 72 samples, and high-grade squamous intraepithelial lesions (HSIL) in 50 samples. Visual inspection with acetic acid (VIA), the standard for cervical lesion detection in Tanzania, was carried out immediately following sample collection. Twenty-six patients were found to be VIA positive for cervical lesions and 115 were VIA negative. All VIA-positive samples were identified as LSIL or HSIL, while several VIA-negative samples were found to be NILM, LSIL, or HSIL by pap smear. Odds ratios were used to identify risk factors for testing VIA positive. Testing HIV+, HSIL, having >5



**FIG 1** Effect of HPV status upon the cervical microbiome diversity. (A) A total of 134 cohort DNA samples were genotyped for HIV and HPV status. The frequency of samples were graphed as the following groups: HIV- HPV- ( $n = 8$ ), HIV- HPV+ ( $n = 87$ ), HIV+ HPV- ( $n = 0$ ), and HIV+ HPV+ ( $n = 39$ ). Taxonomic groups were determined by analysis of 16S deep sequencing results of bacterial DNAs. (B) Alpha diversity is graphed as a function of HIV- HPV- ( $n = 8$ ) and HIV- HPV+ ( $n = 87$ ). A  $t$  test showed a significant difference between the HPV- and HPV+ groups ( $P = 0.02$ ). (C) Bacterial diversity is graphed with each phylum represented as a different color. The color code representing each bacterial phylum is shown in the legend to the right.

sexual partners, and having been infected with a sexually transmitted infection (STI) were identified as significant risk factors for positive VIA status ( $P = 0.0001$ ,  $P = 0.038$ ,  $P = 0.006$ , and  $P = 0.0008$ , respectively).

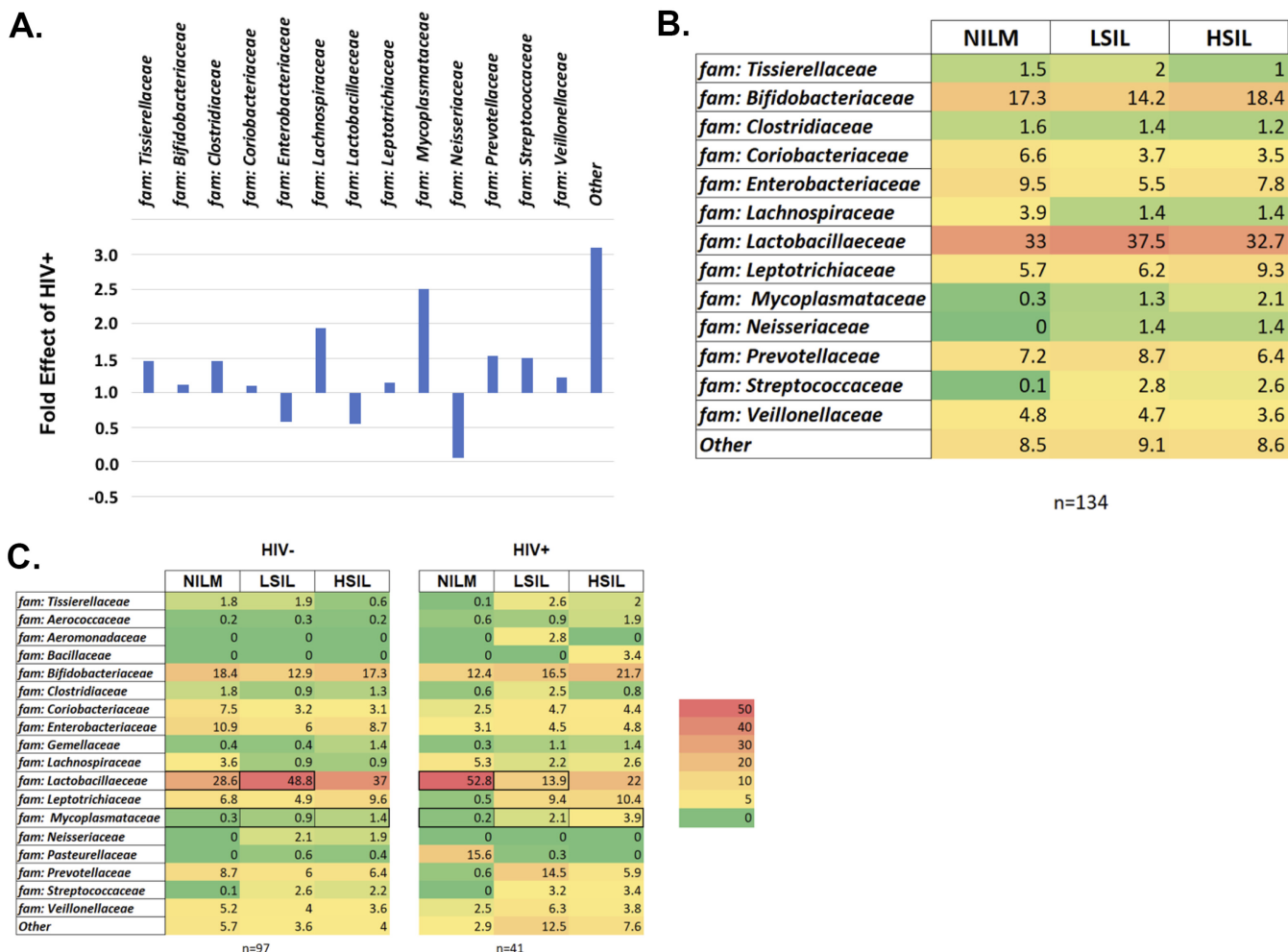
**Cervical bacteria composition and richness.** Samples rarefied to an even depth (1,000 reads) were used to generate 813 operational taxonomic units (OTUs). To assess whether the sampling depth was adequate, rarefaction curves were generated using observed OTUs for HIV status and cervical cytology (see Fig. S1 in the supplemental material). Rarefaction curves for both did not converge but showed a diminishing rate of new OTU identification as the number of reads per sample increased, implying that sampling depth was adequate for evaluating dominant members of the cervical bacterial community. Good's coverage test showed that the sequencing depth was able to characterize 99.4% of the bacterial community on average.



**FIG 2** Phylum-level taxonomy of the cervical bacterial community composition as a function of HIV status and cervical cytology. (A) Phylum-level bacterial taxonomy of the cohort is displayed by HIV status. (B) Phylum-level bacterial taxonomy of the cohort is displayed as a function of cervical cytology. (C) Phylum-level taxonomy of HIV-negative patients as a function of cervical cytology grade. (D) Phylum-level taxonomy of HIV-positive patients as a function of cervical cytology grade. Each phylum is represented as a percentage of the total.

The taxonomic analysis of the reads revealed the presence of six main phyla (relative abundance of >1%) in the cervical epithelium, regardless of HIV or cervical cytology status (Fig. 2). *Firmicutes* was the predominant phylum across all sampling groups, accounting for 41.3% of total reads. The average relative abundance of *Firmicutes* decreased slightly in HIV+ samples compared to HIV- samples (44.4% to 40.2%) and varied by cervical cytology, though no obvious trend was apparent. When considering only the HIV+ samples, the relative abundance of *Firmicutes* appeared to decrease in patients with cervical lesions. *Firmicutes* reads were primarily from the genus *Lactobacillus*, which accounted for 21.9% of total reads. *Tenericutes* accounted for 1.5% of total reads and showed a clear increase in relative abundance with increasing severity of cervical lesions. In HIV- patients, *Tenericutes* increased from 0.3% of reads in NILM patients to 1.3% in HSIL patients (Fig. 2C). In HIV+ patients, the shift is larger; the relative abundance of *Tenericutes* increased from 0.2% in NILM patients to 5.0% in HSIL patients (Fig. 2D). *Tenericutes* reads were primarily assigned to the *Mycoplasma* and *Ureaplasma* genera, which account for 1.1% and 0.2% of total reads, respectively. *Proteobacteria*, *Fusobacteria*, *Bacteroidetes*, and *Actinobacteria* had smaller or less consistent shifts in relative abundance between HIV and cervical cytology categories. The relative abundance of *Tenericutes* and *Bacteroidetes* were significantly different between HIV+ and HIV- groups ( $P = 0.020$  and  $P = 0.017$ , respectively). No other phyla reached significance on the basis of HIV status or cervical cytology. Comparison of the relative abundance of bacterial families (Fig. 3) found that *Mycoplasmataceae* and *Prevotellaceae* were significantly more abundant in HIV+ patients ( $P = 0.03$  and  $P = 0.07$ , respectively). No families were found to be significantly different in abundance on the basis of cervical cytology alone. However, when analyzed among HIV+ patients, *Prevotellaceae* was found to be significantly more abundant in cervical lesions ( $P = 0.068$ ).

**Cervical bacterial diversity estimates.** Alpha diversity metrics, Chao1, observed OTUs, and PD Whole Tree, displayed higher ( $P = 0.009$ ) bacterial richness in HIV+

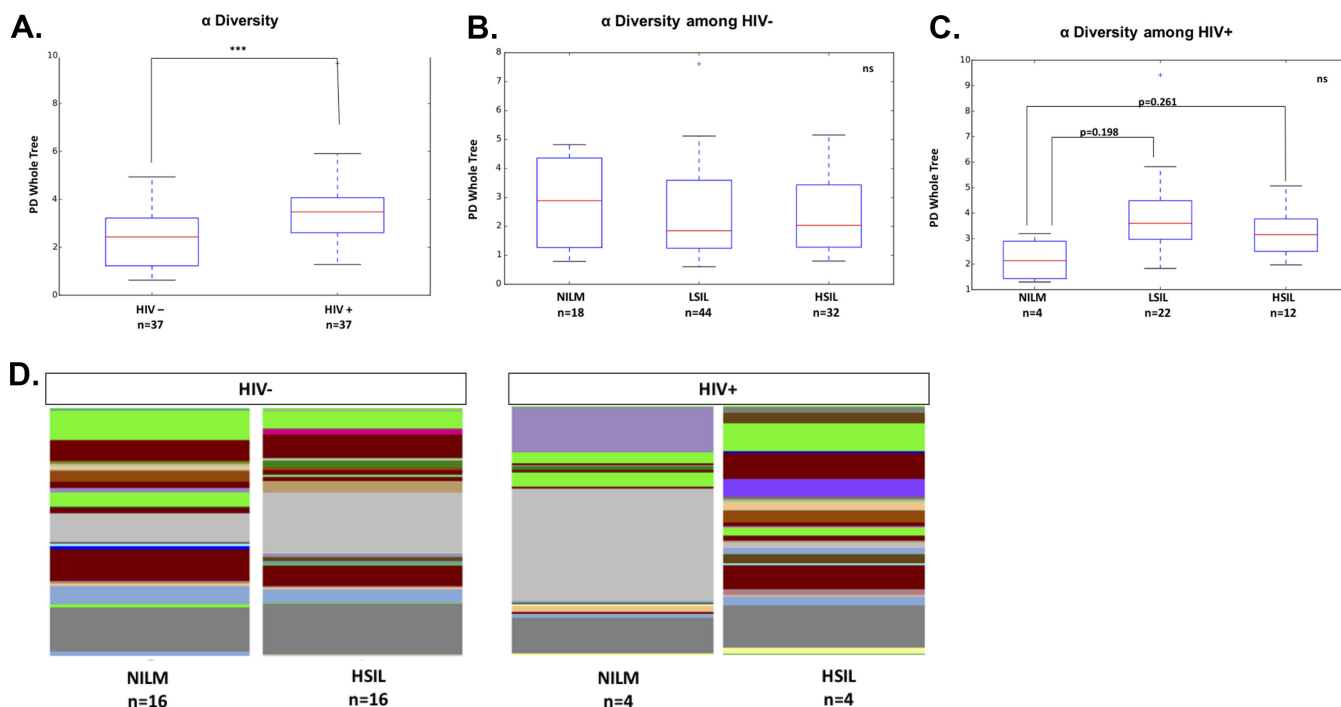


**FIG 3** Relative abundance heatmap of family-level taxonomy of the cervical bacterial community composition as a function of HIV status and cervical cytology. (A) Fold effect of HIV+ on the family-level bacterial taxonomy within the cohort (normalized to 1). (B) Relative abundance heatmap of the family-level taxonomy of cohort versus cervical cytology. (C) Relative abundance heatmap of the family-level taxonomy of the cohort by cervical cytology, separated by HIV status. The data are presented as percentages of the total. The scale is shown to the right of the heatmap.

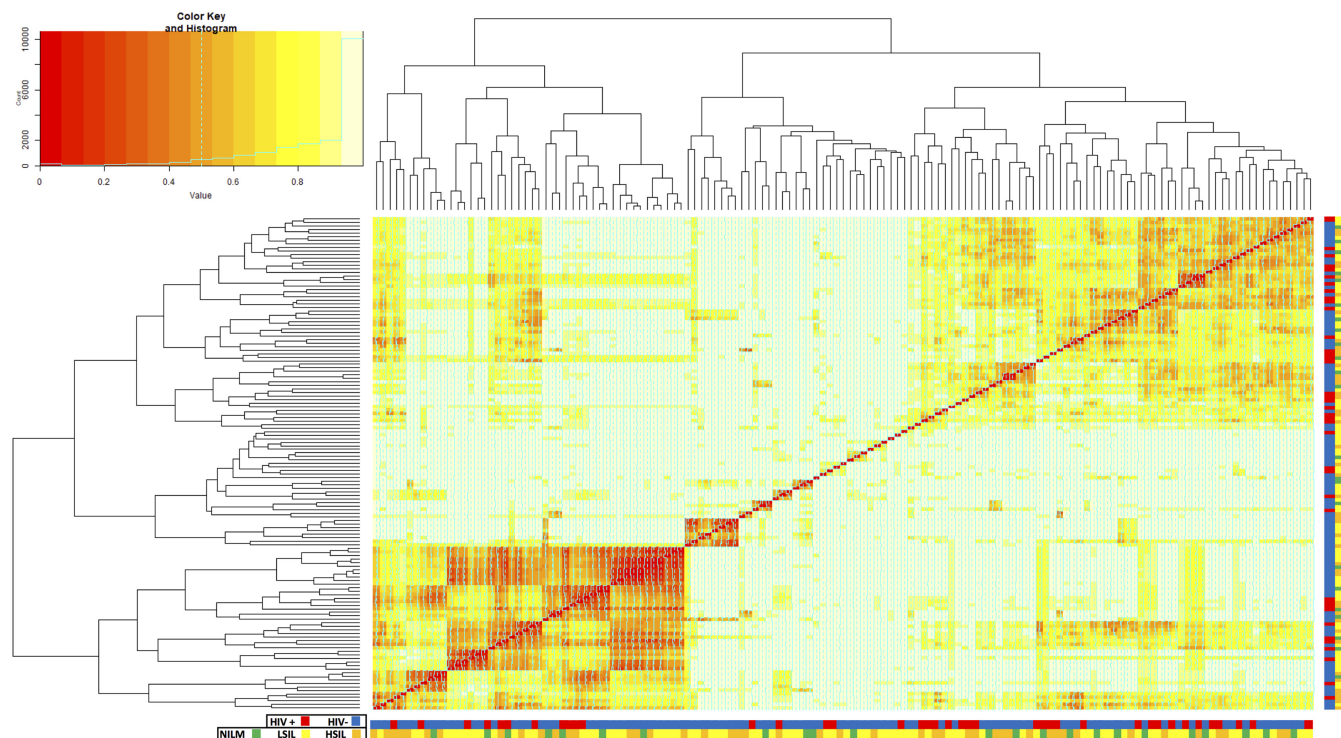
patients than in HIV- patients (Fig. 4). A subset of these samples was matched such that the HIV- and HIV+ groups consisted of the same number of samples, with the same average age, and the same contribution of each cervical cytology to help to control for effects of these confounding variables and to ensure that differences in diversity estimates are not due to differences in sample size. In this matched subset, estimates also displayed higher ( $P = 0.003$ ) bacterial richness in HIV+ patients.

Alpha diversity metrics were similar ( $P > 0.50$ ) for the samples from patients at different cervical cytology grades (NILM, LSIL, or HSIL) in both matched and unmatched sets. When alpha diversity metrics were compared between cervical cytology groups separately for HIV+ samples, LSIL and HSIL trended toward a higher diversity compared to NILM ( $P = 0.198$  and  $P = 0.261$ , respectively). Analysis of age-matched, HIV+ NILM/HSIL pairs maintained this trend ( $P = 0.264$ ; Chao1  $P = 0.13$ ). Comparison of the relative abundance of genus-level reads between these groups showed a noticeably more diverse profile for HSIL samples, which lack the dominance of *Lactobacillus* and *Haemophilus* seen in NILM samples.

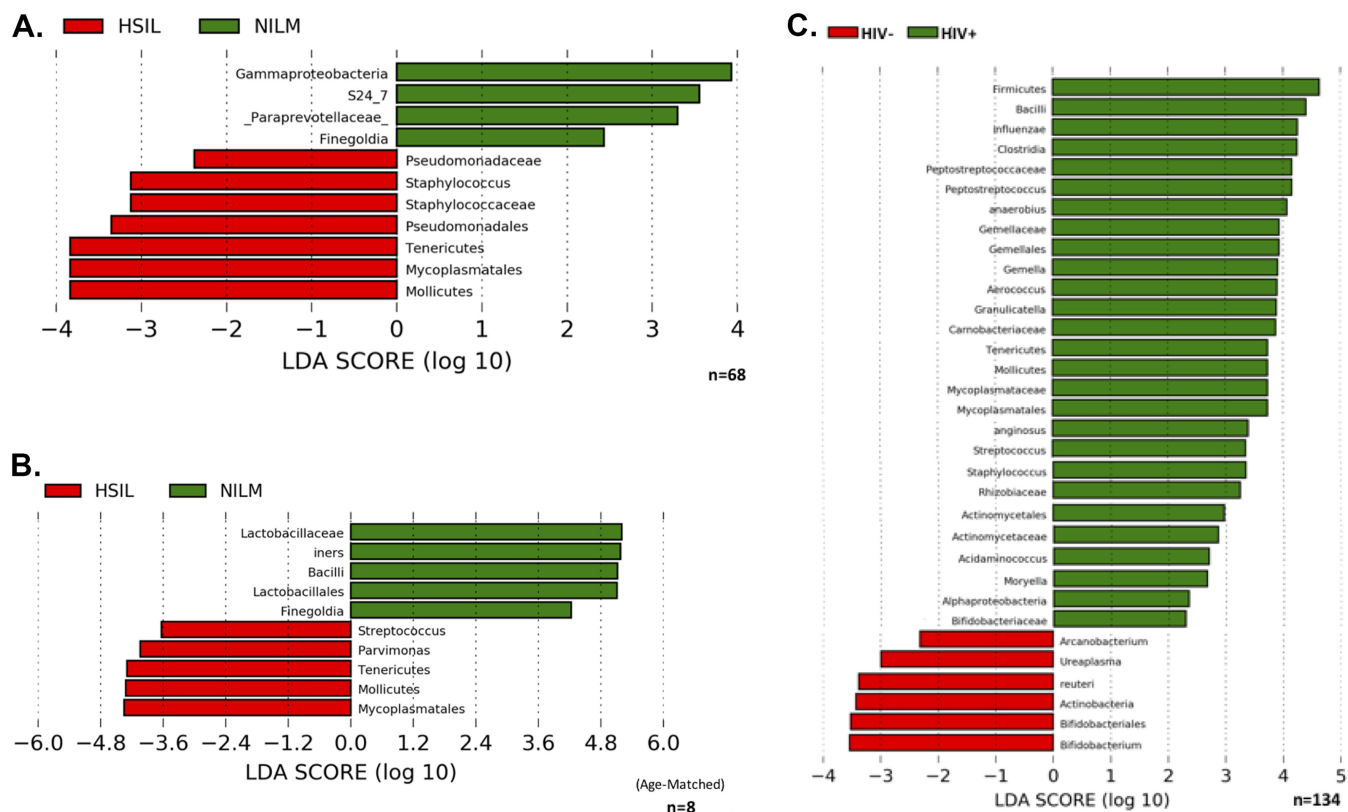
Beta diversity analysis showed that bacterial communities were quite varied between samples (Fig. 5); no discrete communities characterized a large number of samples. On average, the cervical bacterial communities of HIV-positive patients were shown to be significantly different from the communities of HIV-negative patients ( $P =$



**FIG 4** Alpha diversity measurements of cohort subgroups. (A) Relationship between HIV status and alpha diversity of cervical bacteria. (B) Relationship between cervical cytology and bacterial alpha diversity in HIV- individuals. (C) Relative abundance of genus-level reads differentiated by cervical cytology in HIV+ and HIV- individuals. Statistical significance is indicated as follows: ns, not significant; \*,  $P < 0.1$ ; \*\*,  $P < 0.05$ ; \*\*\*,  $P < 0.01$ . (D) Relative bacterial diversity of cervical microbiota graphed as a function of HIV status. Each color represents a different taxonomic family as defined by deep sequencing of the 16S gene.



**FIG 5** Heatmap of the Bray-Curtis distances between each sample (beta diversity). Samples are grouped into a similarity tree based on the abundance of each OTU. Lower values (red) indicate more similarity. HIV status and cervical cytology of each sample are indicated by color beneath each column and beside each row (HIV+ [red], HIV- [blue], NILM [green], LSIL [yellow], HSIL [orange]).



**FIG 6** LefSe linear discriminant analysis (LDA) scores. Microbes associated with cervical cytology status and/or HIV status are displayed. (A) Taxonomies differentiating bacterial microbiota in cytologically normal versus HSIL cervixes. (B) Taxonomies differentiating bacterial microbiota in cytologically normal versus HSIL cervixes in age-matched HIV+ patients. (C) Taxonomies differentiating bacterial microbiota in HIV- versus HIV+ cervixes.

0.001). Similarly, patients who tested positive for HPV tended to have different bacterial communities from those who tested negative for HPV ( $P = 0.008$ ). Bacterial communities were also shown to differ significantly depending on cervical cytology among HIV-positive patients ( $P = 0.05$ ).

**Bacteria associated with cervical cytology states and/or HIV status.** Linear discriminant analysis effect size (LEfSe) was used to identify bacterial taxonomies which differentiate cervical microbiota in normal individuals (NILM) from microbiota in patients with precancerous lesions (HSIL). The sum of reads at each taxonomic rank was considered. *Gammaproteobacteria*, *s24\_7*, *Paraprevotellaceae* (nonverified taxonomy), and *Finiegoldia* associated with NILM cervixes, while *Pseudomonadaceae*, *Staphylococcus*, and *Mycoplasmatales* associated with precancerous lesions. *Mycoplasmatales* were dominant among *Tenericutes*, resulting in the significant association seen between the phylum and cervical lesions. A distance-based redundancy analysis (db-RDA) analysis of bacterial communities as a function of HIV and/or cervical cytology is summarized in Fig. S2. LefSe was then used to compare HIV+, age-matched pairs of NILM and HSIL patients to determine which bacteria may influence the development of lesions in high-risk, HIV+ populations. *Mycoplasmatales* were most strongly associated with cervical lesions in HIV+ patients, followed by *Parvimonas* and *Streptococcus*. In NILM patients, an abundance of *Lactobacillus*, especially *Lactobacillus iners* was found, and somewhat less significantly *Finiegoldia*. LefSe analysis of samples by HIV found several bacteria to be associated with being HIV+ (Fig. 6C). An abundance of non-*Lactobacillus* bacilli was the most significant differentiating taxonomy between HIV-positive and -negative samples. *Mycoplasma* was also associated with HIV+ individuals, supporting the significant difference in relative abundance between HIV-positive and -negative groups shown previously using a direct Kruskal-Wallis comparison. Interestingly, *Ureaplasma* (a member of *Mycoplasmatales*) and *Lactobacillus reuteri* were associated with

HIV– patients, while other members of their respective families were associated with HIV+ patients. This suggests the existence of metabolic niches in the cervical microbiome which may be populated by pathogenic or nonpathogenic bacteria.

## DISCUSSION

We found that HPV was in high abundance in the cohort (Fig. 1A). All of the patients who were HIV positive were positive for one or more HPVs (Fig. 1A). Among the HIV– samples, HPV was associated with at least a 10-fold increase in *Bacteroidetes* and fusobacteria as well as a decrease in *Actinobacteria* (Fig. 1B and C). Previous studies support the conclusion that HPV affects the microbiome (24).

Certain members of the cervicovaginal microbiome are known to protect against infection and pathogenesis. The primary defense mechanisms of the cervicovaginal mucosa are antimicrobial peptides, a pH of less than 4.5, and a microbiome dominated by lactobacilli. An imbalance in these defenses can result in physiochemical changes that produce alterations of the vaginal mucosa and cervical epithelium (27). In particular, an abundance of *Lactobacillus crispatus* shows an inverse relationship with detectable or symptomatic HIV, HPV, or herpesvirus infection (25). This suggests that other cervicovaginal microbes may be important in preventing or enhancing the acquisition and pathogenesis of such infections. Microbes that are associated with enhanced pathogenesis have largely gone unidentified or unstudied, especially in the population most at risk, HIV-positive women in sub-Saharan Africa.

In this study, HIV was shown to have a significant effect on the cervical microbiome, increasing bacterial richness and decreasing beta diversity. These results are similar to what has been reported for the cervicovaginal microbiome and suggest that changes in the cervical epithelium microenvironment brought on by HIV exert some selective pressure on cervical bacterial communities (24–28). *Mycoplasma* was significantly more abundant in HIV-positive patients and was found to be one of the main categories of bacteria that differentiate the cervical microbiota of HIV-positive and HIV-negative individuals (Fig. 3). Interestingly, bacteria of the order *Bacilli*, of which *Lactobacillus* is a member, were strongly associated with HIV-positive patients. The absence of *Bacilli* reads classified as *Lactobacillus* among the significant factors of HIV-positive cervical microbiota suggests that this may be due to a shift from protective to nonprotective *Bacilli* in HIV+ individuals. When the cohort was analyzed without taking HIV status into account, cervical cytology did not appear to have a statistically significant association with differences in the cervical microbiome (Fig. 4B). However, when HIV was controlled for by separating analysis by groups of HIV-positive or HIV-negative patients only, differences in cervical bacterial communities that varied on the basis of cervical lesion status began to reach statistical significance (Fig. 4C). This suggests that development of precancerous cervical lesions is associated with a certain microbiota. Among these microbiota, *Mycoplasmatales* stood out as the most significant differentiator between the cervical microbiota of a cervix with precancerous lesions from a cervix without precancerous lesions (Fig. 6). Bacteria belonging to the order *Mycoplasmatales* also showed the clearest linear increase in abundance with development of more severe lesions in both HIV-positive and HIV-negative populations. The most common *Mycoplasmatales* to infect the urogenital tract of women are *Mycoplasma genitalium* and *Mycoplasma hominis*. *M. genitalium* and *M. hominis* are noncommensal bacteria commonly associated with cervicitis, BV, PID, and HIV infection, though *M. genitalium* has been much better studied (29–31). It is not well understood whether HIV promotes mycoplasma infection or persistence of an otherwise transient infection in an HIV-negative individual. One study found that HIV-positive women cleared *M. genitalium* infections more slowly than HIV-negative women did, and the infection recurred in 39% of the patients after clearance (32). The role of *M. genitalium* infection in influencing initial infection of HIV also remains unclear; however, a strong association between the severity of *M. genitalium* infection and HIV shedding from the cervix has been shown (33). What is clear is that *M. genitalium* infects the epithelia, disrupting tight junctions, and inducing a chronic inflammatory response. The potential for *M. genitalium* to



influence replication of HIV suggests that host innate responses to *M. genitalium* infection may influence pathogenesis of other sexually transmitted infections. Induction of HPV in this way is particularly interesting based on the association between *Mycoplasma* and cervical lesions. Infection with *M. genitalium* increases the rate of infection with an HPV genotype associated with a high risk of developing cervical cancer (34). Recent work has shown that mycoplasma also increases the risk of development of cervical lesions, supporting the association we report in this study (34). *Mycoplasma* can establish persistent, intracellular infections in epithelial cells, which may lead to bacterial vaginosis and/or cervicitis. *M. genitalium* has been established as an independent, causal microbe responsible for cervicitis (35). This suggests that *Mycoplasma* may act as both an intracellular and extracellular stressor, particularly if coinfection with HPV has taken place. This interaction would most likely involve inflammatory cytokines induced by *Mycoplasma* infection. Further study is needed to determine whether the inflammatory cytokines induced by *Mycoplasma* infection include cytokines that are associated with precancerous cervical lesions.

*Mycoplasma* is a low-abundance microbe that has been shown to cause cervicitis. However, the lack of significant associations in previous metagenomic studies is largely due to a lack of optimization of statistical analyses for the presence of low-abundance microbes. In our study, *Mycoplasma* was a prominent result, likely due to the large HIV-positive proportion of the cohort, wherein immunosuppression allowed higher abundance of the bacteria to accumulate. There was a linear increase in the abundance of *Mycoplasmatales* from NILM to HSIL seen in both HIV-positive and -negative groups.

In this study, we took great effort to control for variation in the cervical microbiome so as to reduce confounding effects that might obscure the bacterial communities that were associated with HPV pathogenesis. The HIV-positive population is of particular interest, since they appear to show changed cervical microbiota associated with HPV pathogenesis (Fig. 2, 3, 4, and 6). Future studies, recruiting a cohort of all HIV-positive women with and without cervical lesions would be desirable in order to better characterize HIV-associated microbiota which promote HPV infection and progression to cervical cancer. Currently, few cervical microbiome cohort studies have been conducted in HIV-positive populations. It is clear that variables such as diet, genetic background, antibiotics or ART, can dramatically affect the microbiota and thus should be carefully controlled at the point of recruitment to the study.

Longitudinal studies of the cervical microbiome are needed to understand how microbe populations change over time, particularly in individuals with HSIL. Long-term longitudinal studies will allow determination of early changes in the cervical microbiota that may help predict the development of precancerous lesions. Because progression of HPV infection to cervical cancer is a process that takes decades, and in many individuals never reaches cancer at all, such a study would need to be large. Studies of the cervical microbiome can be further improved using metagenomic sequencing, rather than 16S or other targeted sequencing techniques that lack depth. 16S amplification ignores microbes that lack a gene to match the primers, for example, viruses, archaea, and eukaryotes are not accounted for. Because only a portion of one gene is being sequenced, the microbes present may be estimated only to the genus level or to a higher taxonomic level. Since the majority of medium- or large-scale cervicovaginal microbiome studies have used this method, the role of nonbacterial components of cervicovaginal microbiome in HPV infection and disease has not been characterized.

As the world's HIV-positive population grows, cervical cancer is expected to become an even more significant problem, despite increasing coverage of antiretroviral treatment (ART). Compared to the risk reduction after ART seen in other AIDS-defining cancers like Kaposi's sarcoma and non-Hodgkin's lymphoma, the risk of cervical cancer is not significantly affected, and recurrence rates remain high with or without treatment (36–39). Understanding microbes that influence this environment will help identify cervical microbiota associated with low- and high-grade cervical lesions. This may allow certain cervical microbiota to be used as diagnostic markers for those at high risk of developing cervical cancer and for the development of preventative probiotic or

antibiotic treatments that could control the cervical microbiome by promoting bacterial colonization with a microbiota associated with healthy cervical cytology. Our studies have identified a unique microbiota associated with HSIL. Data derived from our precise sampling of cervical lesions lead us to propose that *Mycoplasma* contributes to a cervical microbiome status that promotes HPV-related cervical lesions. These results suggest a greater influence of the bacterial microbiota on the outcome of HPV infection than previously thought.

## MATERIALS AND METHODS

**Participants and ethical precautions.** This study reports findings derived from a larger cross-sectional cohort study analyzing demographics of HPV and cervical cancer in HIV-positive and -negative women from rural and urban Tanzania.

The cervical microbiome study participants were part of a larger ongoing study to follow HIV- and HPV-associated cervical dysplasia in women at Ocean Road Cancer Institute (ORCI), the only cancer treatment hospital in Tanzania. Between March 2015 and February 2016, female patients undergoing cervical cancer screening were approached for enrollment in the study. Those who were pregnant, menstruating, under 18, reported being sick in the past 30 days, or had a preexisting, non-HIV, immunologic defect were excluded from the study. Disease histories as well as physical examinations were carried out to rule out any clinical symptoms or visible signs for these conditions. Samples were collected at three sites in Tanzania: ORCI in Dar es Salaam and rural clinics in Chalinze and Bagamoyo. A total of 144 cervical cytobrush samples obtained from these women were sequenced, of which 134 samples produced at least 1,000 reads and complete demographic data was available for the women. Of these, 132 had complete HIV data and cervical cytology reads.

**Demographic data collection.** All study participants gave informed consent and were evaluated by study clinicians. A set of pretested, standardized questionnaires was used to gather demographic data. All personal identifiers were removed from samples to ensure patient confidentiality. With the permission of the patients, medical history was retrospectively retrieved from hospital medical records. More than 30 variables were identified and assessed in the questionnaire. The current study uses only data collected regarding age and laboratory test results (pap smears, visual inspection with acetic acid [VIA], CD4 count, genotyping of HPV, results of serological testing for HIV-1).

**Specimen collection, HIV, CD4, and pap tests.** Blood samples were collected via venipuncture into acid-citrate-dextrose tubes and processed using centrifugation at the on-site study laboratory within 6 h of being drawn. The separated plasma was tested at the ORCI, as part of standard of care, using Standard Diagnostics HIV-1/2 3.0 detection kit and BD products CD4 FITC, CD8 PE, and CD3 Per CP antibodies to test the CD4 counts using a BD Accuri C6 Plus. Cervical cytobrush samples and pap smears were collected from all patients. Pap smears were examined by at least three trained cytologists and classified according to the pap classification protocol: negative for intraepithelial lesion or malignancy (NILM); atypical squamous cells of undetermined significance (ASC-US); low-grade squamous intraepithelial lesions (LSIL); atypical squamous cells but cannot exclude high-grade lesions (ASC-H); high-grade squamous intraepithelial lesions (HSIL). Cervical cytobrush specimens were placed in lysis buffer and then shipped to the Nebraska Center for Virology at the University of Nebraska-Lincoln (UNL) for processing.

**DNA isolation, 16S rRNA library preparation, and sequencing of the V4 region.** Cervical cytobrush samples were vortexed and separated from the brush with lysis buffer. DNA was extracted from the lysis buffer using the Qiagen Tissue extraction kit (Dneasy) according to the manufacturer's protocol. The DNA concentration was determined by UV spectrophotometer at 260/280 nm.

DNA was then used for tag sequencing of the V4 hypervariable region of the 16S rRNA gene. A 250-bp section of the V4 region was amplified using universal primers described in reference 40. The PCRs were performed in 25  $\mu$ l. The cycling conditions were as follows: an initial denaturation of 98°C for 3 min, followed by 25 cycles, with 1 cycle consisting of denaturation at 98°C for 30 s, annealing at 55°C for 30 s, and extension at 68°C for 45 s, and then a final elongation of 68°C for 4 min. Following amplification, PCR products were analyzed on a 2% agarose gel to confirm correct product size. Normalized amplicons (1 to 2 ng/ $\mu$ l) from 144 samples were pooled together using an epMotion M5073 liquid handler (Eppendorf AG, Hamburg, Germany). Pooled libraries were sequenced using the Illumina MiSeq platform using the dual-index sequencing strategy outlined by Kozich et al. (40).

**HPV genotyping.** To determine HPV status, DNA samples were subjected to HPV redundant primer using the GP5+/GP6+ primer set, which detect up to 40 different mucosal HPVs (41–43). Samples found to be HPV positive were genotyped for HR-HPVs (types 16, 18, 30, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 66) and LR-HPVs (types 6 and 11) using a low-cost multiplex PCR assay (44).

**Data processing and bacterial community analysis.** The sequencing data obtained from the sequencer was subsequently analyzed using the Illumina MiSeq data analysis pipeline developed by the Fernando lab (described in detail at <https://github.com/FernandoLab>). Briefly, initial quality filtering was carried out to remove sequences that had ambiguous bases, incorrect lengths, and inaccurate assemblies. Subsequently, the quality-filtered reads were run through the UPARSE pipeline (<http://www.drive5.com/uparse/>) and subjected to chimera filtering and OTU clustering (at a similarity threshold of 97%), followed by the generation of an OTU table. Taxonomy was assigned to the OTUs using the `assign_taxonomy.py` command available in QIIME using the latest version of the Greengenes database (May 2013).

**Statistical analyses.** The OTU table was rarefied across samples to the lowest sample depth (1,000 reads) using QIIME based on the Mersenne Twister pseudorandom number generator. All statistical

analyses were performed with samples at an even depth. Bar charts summarizing average taxonomic makeup of samples by HIV status and cervical cytology were constructed from the rarefied OTU table in QIIME. Heatmaps showing the relative abundance of bacterial taxonomic families were constructed using the "plot\_ts\_heatmap" command using the mctoolsR package for R. Differences in bacterial families by HIV status or cervical cytology were evaluated using the "taxa\_summary\_by\_sample\_type" command in mctoolsR using Kruskal-Wallis. Families with less than 1% abundance were excluded in this analysis. Alpha diversity estimators Chao1, observed OTUs, and PD whole tree and rarefaction curves were calculated for the overall bacterial community using QIIME. Good's coverage test was performed to evaluate whether adequate sampling depth was achieved. Mean alpha diversity estimates for HIV-positive, HIV-negative, NILM, LSIL, and HSIL groups were compared using nonparametric two-sample *t* tests using Monte Carlo permutations in QIIME. The weighted and unweighted UniFrac distance matrix for bacterial communities were calculated using QIIME. Even depth across samples avoided biases that could be encountered when using the UniFrac metric (45). Bacterial community composition differences were evaluated using the unweighted UniFrac distance matrix as an input for a distance-based redundancy analysis (db-RDA) in QIIME, where HIV status, cervical cytology, and HPV status were used as main effects. A heatmap was generated using the heatmap.2 command in the "ggplots" package for "R" using the Bray-Curtis distance matrix to visualize relationships between samples. Significance was declared at  $P \leq 0.1$  throughout this study. The linear discriminant analysis effect size (LEfSe) was used to identify specific OTUs that differed HIV status and cervical cytology (46). LEfSe uses a nonparametric factorial Kruskal-Wallis rank sum test followed by a linear discriminant analysis to identify both statistically significant and biologically relevant features. The relative abundances of the OTUs were used as input for LEfSe. Demographic data were examined using odds ratio and an associated *P* value to test for factors associated with HIV status and/or a positive VIA status. All *P* values are reported as FDR-corrected *P* values.

**Ethics statement.** All human subject protocols were approved by safety committees at the Ocean Road Cancer Institute (ORCI) and UNL in accordance with the Helsinki Declaration. Participation by patients was entirely voluntary, and written patient consent was required for inclusion in the study.

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mBio.02785-18>.

**FIG S1**, TIF file, 1.1 MB.

**FIG S2**, TIF file, 1.1 MB.

## ACKNOWLEDGMENTS

We thank the Angeletti and Wood laboratory members and members of the National Center for Virology (NCV) for critical discussions of this work. We thank members of the ORCI and the Tanzanian Ministry of Health, who facilitated these studies. We thank Danielle Shea for logistical and technical support related to this project.

## REFERENCES

- Forman D, de Martel C, Lacey CJ, Soerjomataram I, Lortet-Tieulent J, Bruni L, Vignat J, Ferlay J, Bray F, Plummer M, Franceschi S. 2012. Global burden of human papillomavirus and related diseases. *Vaccine* 30(Suppl 5):F12–F23. <https://doi.org/10.1016/j.vaccine.2012.07.055>.
- De Vuyst H, Alemany L, Lacey C, Chibweshu CJ, Sahasrabudde V, Banura C, Denny L, Parham GP. 2013. The burden of human papillomavirus infections and related diseases in sub-Saharan Africa. *Vaccine* 31:F32–F46. <https://doi.org/10.1016/j.vaccine.2012.07.092>.
- Adler DH, Wallace M, Bennie T, Mrubata M, Abar B, Meiring TL, Williamson A-L, Bekker L-G. 2014. Cervical dysplasia and high-risk human papillomavirus infections among HIV-infected and HIV-uninfected adolescent females in South Africa. *Infect Dis Obstet Gynecol* 2014:498048. <https://doi.org/10.1155/2014/498048>.
- Salazar KL, Zhou HS, Xu J, Peterson LE, Schwartz MR, Mody DR, Ge Y. 2015. Multiple human papilloma virus infections and their impact on the development of high-risk cervical lesions. *Acta Cytol* 59:391–398. <https://doi.org/10.1159/000442512>.
- Williamson A-L. 2015. The interaction between human immunodeficiency virus and human papillomaviruses in heterosexuals in Africa. *J Clin Med* 4:579–592. <https://doi.org/10.3390/jcm4040579>.
- McDonald AC, Tergas AI, Kuhn L, Denny L, Wright TC, Jr. 2014. Distribution of human papillomavirus genotypes among HIV-positive and HIV-negative women in Cape Town, South Africa. *Front Oncol* 4:48. <https://doi.org/10.3389/fonc.2014.00048>.
- Memiah P, Makokha V, Mbuthia W, Kiiru GW, Agbor S, Odhiambo F, Ojoo S, Mbizo J, Muhula S, Mahasi G, Biadgilign S. 2015. Epidemiology of cervical squamous intraepithelial lesions in HIV infected women in Kenya: a cross-sectional study. *Afr J Reprod Health* 19:133–139.
- Menon SS, Rossi R, Harebottle R, Mabeya H, Vanden Broeck D. 2016. Distribution of human papillomaviruses and bacterial vaginosis in HIV positive women with abnormal cytology in Mombasa, Kenya. *Infect Agent Cancer* 11:17. <https://doi.org/10.1186/s13027-016-0061-1>.
- Ezechi OC, Ostergren PO, Nwaokorie FO, Ujah IA, Odberg Pettersson K. 2014. The burden, distribution and risk factors for cervical oncogenic human papilloma virus infection in HIV positive Nigerian women. *Virolgy J* 11:5. <https://doi.org/10.1186/1743-422X-11-5>.
- Hanisch RA, Sow PS, Toure M, Dem A, Dembele B, Toure P, Winer RL, Hughes JP, Gottlieb GS, Feng Q, Kiviat NB, Hawes SE. 2013. Influence of HIV-1 and/or HIV-2 infection and CD4 count on cervical HPV DNA detection in women from Senegal, West Africa. *J Clin Virol* 58:696–702. <https://doi.org/10.1016/j.jcv.2013.10.012>.
- Mitra A, MacIntyre DA, Marchesi JR, Lee YS, Bennett PR, Kyrgiou M. 2016. The vaginal microbiota, human papillomavirus infection and cervical intraepithelial neoplasia: what do we know and where are we going next? *Microbiome* 4:58. <https://doi.org/10.1186/s40168-016-0203-0>.
- Kyrgiou M, Mitra A, Moscicki AB. 2017. Does the vaginal microbiota play a role in the development of cervical cancer? *Transl Res* 179:168–182. <https://doi.org/10.1016/j.trsl.2016.07.004>.
- Guijon F, Paraskevas M, Rand F, Heywood E, Brunham R, McNicol P. 1992. Vaginal microbial flora as a cofactor in the pathogenesis of uterine

- cervical intraepithelial neoplasia. *Int J Gynaecol Obstet* 37:185–191. [https://doi.org/10.1016/0020-7292\(92\)90379-W](https://doi.org/10.1016/0020-7292(92)90379-W).
14. Gillet E, Meys JF, Verstraelen H, Bosire C, De Sutter P, Temmerman M, Vanden Broeck D. 2011. Bacterial vaginosis is associated with uterine cervical human papillomavirus infection: a meta-analysis. *BMC Infect Dis* 11:10. <https://doi.org/10.1186/1471-2334-11-10>.
  15. Guo Y-L, You K, Qiao J, Zhao Y-M, Geng L. 2012. Bacterial vaginosis is conducive to the persistence of HPV infection. *Int J STD AIDS* 23: 581–584. <https://doi.org/10.1258/ijsa.2012.011342>.
  16. Vriend HJ, Bogaards JA, van Bergen JEAM, Brink AATP, van den Broek IVF, Hoebe CJP, King AJ, van der Sande MAB, Wolffs PFG, de Melker HE. 2015. Incidence and persistence of carcinogenic genital human papillomavirus infections in young women with or without Chlamydia trachomatis co-infection. *Cancer Med* 4:1589–1598. <https://doi.org/10.1002/cam4.496>.
  17. Clarke MA, Rodriguez AC, Gage JC, Herrero R, Hildesheim A, Wacholder S, Burk R, Schiffman M. 2012. A large, population-based study of age-related associations between vaginal pH and human papillomavirus infection. *BMC Infect Dis* 12:33. <https://doi.org/10.1186/1471-2334-12-33>.
  18. Skapinyecz J, Smid I, Horvath A, Jeney C, Kardos L, Kovacs P. 2003. Pelvic inflammatory disease is a risk factor for cervical cancer. *Eur J Gynaecol Oncol* 24:401–404.
  19. Giraud J, Coiffic J, Poulain P, Kerisit J. 1998. High prevalence of cervical intra-epithelial neoplasia in women treated for pelvic inflammatory disease. *Eur J Obstet Gynecol Reprod Biol* 81:51–54. [https://doi.org/10.1016/S0301-2115\(98\)00146-8](https://doi.org/10.1016/S0301-2115(98)00146-8).
  20. Msuya SE, Mbizvo E, Stray-Pedersen B, Sundby J, Sam NE, Hussain A. 2002. Reproductive tract infections and the risk of HIV among women in Moshi, Tanzania. *Acta Obstet Gynecol Scand* 81:886–893. <https://doi.org/10.1034/j.1600-0412.2002.810916.x>.
  21. Swanepoel PJ, Michelow P, Du Plessis R, Proudfoot IG, Tarr GA, Bockel SL, Swanepoel CJ. 2013. Cervical squamous intraepithelial lesions and associated cervical infections in an HIV-positive population in rural Mpumalanga, South Africa. *Cytopathology* 24:264–271. <https://doi.org/10.1111/j.1365-2303.2012.00998.x>.
  22. Lewis DA. 2011. HIV/sexually transmitted infection epidemiology, management and control in the IUSTI Africa region: focus on sub-Saharan Africa. *Sex Transm Infect* 87(Suppl 2):ii10–ii13. <https://doi.org/10.1136/sextrans-2011-050178>.
  23. Spear GT, Sikaroodi M, Zariffard MR, Landay AL, French AL, Gillevet PM. 2008. Comparison of the diversity of the vaginal microbiota in HIV-infected and HIV-uninfected women with or without bacterial vaginosis. *J Infect Dis* 198:1131–1140. <https://doi.org/10.1086/591942>.
  24. Lee JE, Lee S, Lee H, Song Y-M, Lee K, Han MJ, Sung J, Ko GP. 2013. Association of the vaginal microbiota with human papillomavirus infection in a Korean twin cohort. *PLoS One* 8:e63514. <https://doi.org/10.1371/journal.pone.0063514>.
  25. Borgdorff H, Tsvitivadze E, Verhelst R, Marzorati M, Jurriaans S, Ndayisaba GF, Schuren FH, van de Wijgert JHHM. 2014. Lactobacillus-dominated cervicovaginal microbiota associated with reduced HIV/STI prevalence and genital HIV viral load in African women. *ISME J* 8:1781–1793. <https://doi.org/10.1038/ismej.2014.26>.
  26. Oh HY, Kim BS, Seo SS, Kong JS, Lee JK, Park SY, Hong KM, Kim HK, Kim MK. 2015. The association of uterine cervical microbiota with an increased risk for cervical intraepithelial neoplasia in Korea. *Clin Microbiol Infect* 21:e674.e1–e674.e9. <https://doi.org/10.1016/j.cmi.2015.02.026>.
  27. Audirac-Chalifour A, Torres-Poveda K, Bahena-Román M, Téllez-Sosa J, Martínez-Barnette J, Cortina-Ceballos B, López-Estrada G, Delgado-Romero K, Burguete-García AI, Cantú D, García-Carrancá A, Madrid-Romera V. 2016. Cervical microbiome and cytokine profile at various stages of cervical cancer: a pilot study. *PLoS One* 11:e0153274. <https://doi.org/10.1371/journal.pone.0153274>.
  28. Curty G, Costa RL, Siqueira JD, Meyrelles AI, Machado ES, Soares EA, Soares MA. 2017. Analysis of the cervical microbiome and potential biomarkers from postpartum HIV-positive women displaying cervical intraepithelial lesions. *Sci Rep* 7:17364. <https://doi.org/10.1038/s41598-017-17351-9>.
  29. Napierala Mavedzenge S, Weiss HA. 2009. Association of Mycoplasma genitalium and HIV infection: a systematic review and meta-analysis. *AIDS* 23:611–620. <https://doi.org/10.1097/QAD.0b013e328323da3e>.
  30. Soni S, Alexander S, Verlander N, Saunders P, Richardson D, Fisher M, Ison C. 2010. The prevalence of urethral and rectal Mycoplasma genitalium and its associations in men who have sex with men attending a genitourinary medicine clinic. *Sex Transm Infect* 86:21–24. <https://doi.org/10.1136/sti.2009.038190>.
  31. Irwin KL, Moorman AC, O'Sullivan MJ, Sperling R, Koestler ME, Soto I, Rice R, Brodman M, Yasin S, Droese A, Zhang D, Schwartz DA, Byers RH. 2000. Influence of human immunodeficiency virus infection on pelvic inflammatory disease. *Obstet Gynecol* 95:525–534. <https://doi.org/10.1097/00006250-200004000-00010>.
  32. Vandepitte J, Weiss HA, Kyakuwa N, Nakubulwa S, Muller E, Buvé A, Van der Stuyft P, Hayes R, Grosskurth H. 2013. Natural history of Mycoplasma genitalium infection in a cohort of female sex workers in Kampala, Uganda. *Sex Transm Dis* 40:422–427. <https://doi.org/10.1097/OLQ.0b013e31828bfccf>.
  33. Manhart LE, Mostad SB, Baeten JM, Astete SG, Mandaliya K, Totten PA. 2008. High Mycoplasma genitalium organism burden is associated with shedding of HIV-1 DNA from the cervix. *J Infect Dis* 197:733–736. <https://doi.org/10.1093/infdis/jin103>.
  34. Ye H, Song T, Zeng X, Li L, Hou M, Xi M. 2018. Association between genital mycoplasmas infection and human papillomavirus infection, abnormal cervical cytopathology, and cervical cancer: a systematic review and meta-analysis. *Arch Gynecol Obstet* 297:1377–1387. <https://doi.org/10.1007/s00404-018-4733-5>.
  35. Taylor-Robinson D, Jensen JS. 2011. Mycoplasma genitalium: from chrysalis to multicolored butterfly. *Clin Microbiol Rev* 24:498–514. <https://doi.org/10.1128/CMR.00006-11>.
  36. Cobucci RNO, Lima PH, de Souza PC, Costa VV, Cornetta MDCDM, Fernandes JV, Gonçalves AK. 2015. Assessing the impact of HAART on the incidence of defining and non-defining AIDS cancers among patients with HIV/AIDS: a systematic review. *J Infect Public Health* 8:1–10. <https://doi.org/10.1016/j.jiph.2014.08.003>.
  37. Russomano F, Paz BR, Camargo MJ, Grinstein BG, Friedman RK, Tristao MA. 2013. Recurrence of cervical intraepithelial neoplasia in human immunodeficiency virus-infected women treated by means of electro-surgical excision of the transformation zone (LLETZ) in Rio de Janeiro, Brazil. *Sao Paulo Med J* 131:405–410. <https://doi.org/10.1590/1516-3180.2013.1316578>.
  38. Mungo C, Cohen CR, Maloba M, Bukusi EA, Huchko MJ. 2013. Prevalence, characteristics, and outcomes of HIV-positive women diagnosed with invasive cancer of the cervix in Kenya. *Int J Gynaecol Obstet* 123: 231–235. <https://doi.org/10.1016/j.ijgo.2013.07.010>.
  39. Foulot H, Heard I, Potard V, Costagliola D, Chapron C. 2008. Surgical management of cervical intraepithelial neoplasia in HIV-infected women. *Eur J Obstet Gynecol Reprod Biol* 141:153–157. <https://doi.org/10.1016/j.ejogrb.2008.07.015>.
  40. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 79:5112–5120. <https://doi.org/10.1128/AEM.01043-13>.
  41. Clifford GM, Gallus S, Herrero R, Muñoz N, Snijders PJF, Vaccarella S, Anh PTH, Ferreccio C, Hieu NT, Matos E, Molano M, Rajkumar R, Ronco G, de Sanjosé S, Shin HR, Sukvirach S, Thomas JO, Tunsakul S, Meijer CJLM, Franceschi S. 2005. 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* 366:991–998. [https://doi.org/10.1016/S0140-6736\(05\)67069-9](https://doi.org/10.1016/S0140-6736(05)67069-9).
  42. Ng'andwe C, Lowe JJ, Richards PJ, Hause L, Wood C, Angeletti PC. 2007. The distribution of sexually-transmitted human papillomaviruses in HIV positive and negative patients in Zambia, Africa. *BMC Infect Dis* 7:77. <https://doi.org/10.1186/1471-2334-7-77>.
  43. Chisanga C, Eggert D, Mitchell CD, Wood C, Angeletti PC. 2015. Evidence for placental HPV infection in both HIV positive and negative women. *J Cancer Ther* 6:1276–1289. <https://doi.org/10.4236/jct.2015.615140>.
  44. Samwel K, Kahesa C, Mwaesalage J, Gonzalez D, West JT, Wood C, Palefsky J, Angeletti PC. 2019. Analytical performance of a low-cost multiplex polymerase chain reaction human papillomavirus genotyping assay for use in Sub-Saharan Africa. *J Med Virol* 91:308–316. <https://doi.org/10.1002/jmv.25329>.
  45. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. 2011. UniFrac: an effective distance metric for microbial community comparison. *ISME J* 5:169–172. <https://doi.org/10.1038/ismej.2010.133>.
  46. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. 2011. Metagenomic biomarker discovery and explanation. *Genome Biol* 12:R60. <https://doi.org/10.1186/gb-2011-12-6-r60>.