

# The Cervicovaginal Microbiota and Its Associations With Human Papillomavirus Detection in HIV-Infected and HIV-Uninfected Women

Laura L. Reimers,<sup>1</sup> Supriya D. Mehta,<sup>3</sup> L. Stewart Massad,<sup>6</sup> Robert D. Burk,<sup>1</sup> Xianhong Xie,<sup>1</sup> Jacques Ravel,<sup>7</sup> Mardge H. Cohen,<sup>4</sup> Joel M. Palefsky,<sup>9</sup> Kathleen M. Weber,<sup>5</sup> Xiaonan Xue,<sup>1</sup> Kathryn Anastos,<sup>1</sup> Howard Minkoff,<sup>2</sup> Jessica Atrio,<sup>11</sup> Gypsyamber D'Souza,<sup>8</sup> Qian Ye,<sup>1</sup> Christine Colie,<sup>10</sup> Christine P. Zolnik,<sup>1</sup> Gregory T. Spear,<sup>4,a</sup> and Howard D. Strickler<sup>1,a</sup>

<sup>1</sup>Albert Einstein College of Medicine, Bronx, and <sup>2</sup>Maimonides Medical Center, Brooklyn, New York; <sup>3</sup>University of Illinois–Chicago, <sup>4</sup>Rush Medical Center, and <sup>5</sup>Cook County Health and Hospitals System/Hektoen Institute of Medicine, Chicago, Illinois; <sup>6</sup>Washington University School of Medicine, St. Louis, Missouri; <sup>7</sup>University of Maryland School of Medicine, and <sup>8</sup>Johns Hopkins University, Baltimore, Maryland; <sup>9</sup>University of California–San Francisco; <sup>10</sup>Georgetown University, Washington D.C.; and <sup>11</sup>Montefiore Medical Center, Bronx, New York

**Background.** Bacterial vaginosis (BV) is characterized by low abundance of *Lactobacillus* species, high pH, and immune cell infiltration and has been associated with an increased risk of human papillomavirus (HPV) infection. We molecularly assessed the cervicovaginal microbiota over time in human immunodeficiency virus (HIV)–infected and HIV-uninfected women to more comprehensively study the HPV-microbiota relationship, controlling for immune status.

**Methods.** 16S ribosomal RNA gene amplicon pyrosequencing and HPV DNA testing were conducted annually in serial cervicovaginal lavage specimens obtained over 8–10 years from African American women from Chicago, of whom 22 were HIV uninfected, 22 were HIV infected with a stable CD4<sup>+</sup> T-cell count of > 500 cells/mm<sup>3</sup>, and 20 were HIV infected with progressive immunosuppression. Vaginal pH was serially measured.

**Results.** The relative abundances of *Lactobacillus crispatus* and other *Lactobacillus* species were inversely associated with vaginal pH (all  $P < .001$ ). High (vs low) *L. crispatus* relative abundance was associated with decreased HPV detection (odds ratio, 0.48; 95% confidence interval, .24–.96;  $P_{\text{trend}} = .03$ ) after adjustment for repeated observation and multiple covariates, including pH and study group. However, there were no associations between HPV and the relative abundance of *Lactobacillus* species as a group, nor with *Lactobacillus gasseri*, *Lactobacillus iners*, and *Lactobacillus jensenii* individually.

**Conclusions.** *L. crispatus* may have a beneficial effect on the burden of HPV in both HIV-infected and HIV-uninfected women (independent of pH).

**Keywords.** HIV; HPV; human papillomavirus; microbiota; *L. crispatus*; *Lactobacillus* species.

Human papillomavirus (HPV) infection is the principal etiologic agent in the development of cervical precancer and cancer [1]. Women with human immunodeficiency virus (HIV) infection are at high risk of HPV infection, as well as cervical disease [2], and these risks increase with diminishing CD4<sup>+</sup> T-cell count [3–5]. A prior study by our group also showed that, in both HIV-infected and HIV-uninfected women, the prevalence of HPV is high among those with bacterial vaginosis (BV) [6], a dysbiosis often characterized by a low abundance of *Lactobacillus* species, high pH, and immune cell infiltration. Interestingly, while HPV infection was associated with both BV as well as

HIV in that earlier study, these relationships were independent of one another [6], as BV itself was not more common in HIV-infected than in HIV-uninfected women.

*Lactobacillus* species produce lactic acid, and vaginal pH significantly decreases as the relative abundance (RA) of *Lactobacillus* species (as a group) increases [7]. Some studies have also reported that vaginal bacterial communities specifically dominated by *Lactobacillus crispatus* are associated with lower pH than those dominated by other lactobacilli [8]. This may be important because vaginal pH has been associated with HPV detection. For example, one frequently cited article found a 10%–20% lower risk of HPV detection in subjects with a vaginal pH a < 5.0 than in those with a vaginal pH of  $\geq 5.0$  [9]. It is therefore noteworthy that a recent study found no relation of vaginal pH levels with HIV status [7], a result consistent with our earlier report showing no relation of HIV seropositivity with BV (discussed above).

Overall, while the impact of HIV on BV risk may be limited, BV has been shown to increase local expression and transmission of HIV and might increase risks of certain other sexually transmitted infections (STIs) [6, 10–12]. We hypothesized that

Received 8 April 2016; accepted 5 August 2016; published online 11 August 2016.

Presented in part: 30th International Papillomavirus Conference, Lisbon, Portugal, September 2015. Abstract PS18.

<sup>a</sup>G. T. S. and H. D. S. are senior coauthors.

Correspondence: H. D. Strickler, Department of Epidemiology and Population Health, Albert Einstein College of Medicine, 1300 Morris Park Ave, Belfer Bldg 1312A, Bronx, NY 10461 (howard.strickler@einstein.yu.edu).

The Journal of Infectious Diseases® 2016;214:1361–9

© The Author 2016. Published by Oxford University Press for the Infectious Diseases Society of America. All rights reserved. For permissions, e-mail journals.permissions@oup.com.  
DOI: 10.1093/infdis/jiw374

the commonality that relates both BV and vaginal pH to HPV infection (independent of HIV) is the RA of certain *Lactobacillus* species, most notably *L. crispatus*. Consistent with this, a high RA of *L. crispatus* has been associated with reduced risks of BV and STIs in several studies [8, 12–14]. However, data regarding the cervicovaginal microbiota and its specific relation with HPV infection are limited. In particular, there is a paucity of longitudinal studies of the cervicovaginal microbiota and HPV [15], and even among cross-sectional studies few have involved HIV-infected women. Last, despite considerable speculation that cervicovaginal pH may mediate the effects of the microbiota on HPV infection, no studies have, to our knowledge, assessed whether statistical control for pH weakens the HPV-microbiota association, which would be expected if pH was part of the causal pathway.

Therefore, in the current investigation, we combined data from 2 independently conducted ancillary studies in the Women's Interagency HIV Study (WIHS) cohort: a study involving serial microbiota and vaginal pH testing and a longitudinal study of cervicovaginal HPV detection.

## METHODS

### WIHS Background

The WIHS is an ongoing multiinstitutional cohort investigation of the natural history and pathogenesis of HIV/AIDS in women [16]. Briefly, HIV-infected and at-risk HIV-uninfected women were enrolled through similar sources at 6 sites (Bronx, NY; Brooklyn, NY; Chicago, IL; Los Angeles, CA; San Francisco, CA; and Washington, DC). Enrollment was conducted in 1994–1995 (2054 HIV-infected women and 569 HIV-uninfected women) and again in 2001–2002 (737 HIV-infected women and 406 HIV-uninfected women). Participants undergo a semi-annual clinical visit that involves performance of a Papanicolaou test and collection of a cervicovaginal lavage (CVL) specimen. Written informed consent was obtained from all participants, and the study was approved by each local institutional review board.

### WIHS Substudy Population

As described previously [10], women included in the microbiome substudy were selected from among premenopausal African American participants at the Chicago WIHS site, with at least 8 years of observation and a minimum of 6 annual CVL specimens, with further selection based on their HIV status and CD4<sup>+</sup> T-cell count over time. Specifically, 3 study groups were defined: HIV-uninfected women (n=22); HIV-infected women with rapidly progressive immunosuppression, defined as an initial CD4<sup>+</sup> T-cell count of >500 cells/mm<sup>3</sup> followed by a drop to <200 cells/mm<sup>3</sup>, with a poor immunologic response to therapy (<100 cells/mm<sup>3</sup> increase over a year) during the observation period (n = 20); and HIV-infected women with a stable immune status, defined as a CD4<sup>+</sup> T-cell count

of >500 cells/mm<sup>3</sup> throughout the observation period, who were frequency matched at the index visit to those with rapidly progressive immunosuppression (n = 22). Matching was conducted using MatchIt, based in the R programming language and statistical software [17, 18]. The matching criteria included age (<40 years, 40–49, and >50 years), cigarette smoking, sexual activity (frequency and number of partners), and male condom use. The exclusion criteria were current laboratory detection of gonorrhea, syphilis, trichomoniasis, or chlamydial infection at the index visit or exchange of sex for drugs or money at >5 visits during the observation period. These inclusion/exclusions criteria were intended to limit the factors that needed to be addressed while studying the original scientific question, namely, the influence of HIV on the cervicovaginal microbiota. In a separate independently conducted study, the presence of HPV DNA in CVLs was longitudinally assessed semiannually in all WIHS women with available specimens over a similar period of 8–10 years or more [5].

### Laboratory Testing

#### Vaginal pH and Cervicovaginal Specimens

Semiannually, vaginal pH was measured prior to collection of the CVL by sampling the posterior vaginal pool with a polyethylene terephthalate swab that was then applied to paper strips with a pH range of 4–7 (ColorpHast, EM Reagents, MCB Reagents, Gibbstown, NJ). CVLs were obtained by irrigation of the cervix with 10 mL of nonbacteriostatic sterile saline, followed by aspiration from the posterior fornix. Following collection, CVLs were refrigerated at <10°C until transported to the laboratory on ice within 4 hours, where they were vortexed gently and aliquoted under a hood (using sterile conditions) into 1-mL units and then frozen at –80°C until requested for testing.

#### Detection of HPV DNA

HPV DNA was detected with a L1 consensus primer MY09/MY11/HMB01 polymerase chain reaction (PCR) assay. Primer set PC04/GH20, which amplifies a 268-base pair cellular  $\beta$ -globin DNA fragment, was included as an internal control to assess the adequacy of amplification, as previously described [19–21]. Amplification products were probed for the presence of any HPV DNA with a generic probe mixture and then with filters individually hybridized with type-specific biotinylated oligonucleotide probes for HPV 6, 11, 13, 16, 18< 26, 31, 32, 33, 34, 35, 39, 40, 42, 45, 51, 52, 53, 54, 55, 56, 57, 58, 59, 61, 62, 64, 66, 67, 68, 69, 70, 71, 72, 73, 74, 81, 82, 83, 84, 85, and 89. Oncogenic HPV types were defined according to the recommendations of the International Agency for Research on Cancer [22].

#### Characterization of the Cervicovaginal Microbiota

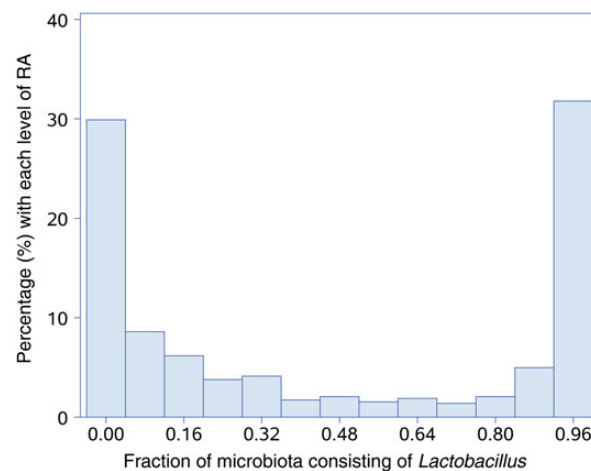
CVL specimens were thawed and centrifuged to pellet microbes. Microbial DNA was isolated using the FastDNA Spin Kit for Soil (MP Biomedicals, Solon, OH). Pyrosequencing of barcoded 16S ribosomal RNA (rRNA) gene V1–V2 region amplicons was

performed as described previously [8, 10]. Briefly, universal primers 27F and 338R were used for PCR amplification of the V1-V2 hypervariable regions of 16S rRNA genes from genomic DNA extracts. The 338R primer included a unique sequence tag to barcode each sample. Barcoded amplicons were pooled in equimolar concentration and pyrosequenced on a Roche/454 Life Sciences FLX instrument. QIIME software [23] was used to bin the sequences based on their barcode, trim the primers and barcodes, and remove sequences with homopolymeric runs longer than 8 base pairs and those with ambiguous base calls and run shorter than 100 base pairs. Detection of chimeric sequences was performed using the UCHIME component of UCLUST, and chimeric sequences were removed. Taxonomic assignments were performed using the Ribosomal Database Project (RDP) Classifier trained on version 10 of the RDP database [24, 25]. The presence of amplification in subject samples was confirmed by gel electrophoresis, while negative controls (no template) were also subjected to PCR and showed no amplification. In general, each sample was run once since our previous studies indicated that duplicates of the PCR amplification and pyrosequencing have low variability [8, 10]. Nonetheless, repeat testing of several current samples across batches and runs was conducted and showed little variability in results between these runs. For the 398 samples in the current study, we obtained 2 387 346 high-quality sequences, with an average of 6266 sequences (range, 183–40 627) from each sample. Despite these several steps and results, as in other microbiome studies, we cannot fully exclude the possibility of bias due to sequencing variability, primer bias, and sequencing platform biases.

### Statistical Methods

We characterized demographic and risk factor data at the index visit for each of the 3 study groups, with statistical differences between groups assessed using analysis of variance (ANOVA; for continuous variables), the Kruskal–Wallis test (for continuous variables with nonnormal distributions), or the  $\chi^2$  test (for categorical data). Correlations between vaginal pH and the RA of *Lactobacillus* species were measured using Spearman correlation. To cluster pyrosequencing 16S rRNA results into corresponding community state types (CSTs), we used hierarchical clustering based on RA with Euclidean distance and Ward linkage. The number of clusters was determined by the Milligan point-biserial criterion [26, 27]. The Shannon diversity index (SDI) was calculated for the microbiome at each patient visit in the data set [28].

The primary outcomes were detection of (1) any HPV and (2) any oncogenic HPV. The associations of SDI or CST with prevalent HPV detection were assessed using generalized estimating equation (GEE) logistic regression, as previously described [29]. These models address serial testing of women over time (ie, repeated cross-sectional measures of association with prevalent HPV detection at each visit) and potential



**Figure 1.** Distribution of relative abundance (RA) of *Lactobacillus* species.

concurrent multitype infection by HPV, while accounting for the intraindividual correlation of the repeated observations. This approach made it possible to incorporate all the available type-specific HPV data at each person-visit, as well as time-updated covariates. It is a highly efficient analytic method, widely used in HPV research [30, 31]. In addition to CSTs, we conducted a priori-planned analyses of the RA of lactobacilli (at the genus and species level), modeled as both continuous and ordinal data. Histograms of *Lactobacillus* species as a group (Figure 1) and individual species suggested similar cut points of low (<7.5%), medium (7.5%–82.5%), and high ( $\geq 82.5\%$ ) RA. Vaginal pH was also examined as a continuous and an ordinal variable (with cut points based on the Amsel criteria for diagnosis of BV [10]). All covariates except study group were time dependent. Furthermore, to study temporal associations related to changes in CSTs and HPV detection over time, we used continuous-time multistate Markov models fit using maximum likelihood, as described in an earlier article [15]. All tests were 2-sided, with a *P* value of .05 used to define significance.

### RESULTS

Select characteristics of the cohort at the index visit are shown in Table 1. The average age of participants was 32.1 years, and by design all subjects were African American and premenopausal (see “Methods” section). Women with rapidly progression of HIV-associated immunosuppression had a lower CD4<sup>+</sup> T-cell count than HIV-infected women with stable immune status and HIV-uninfected women (*P* < .001). A majority of the women in all 3 groups had a lifetime number of male sex partners of >10, were current or former smokers, and had not used oral contraceptives in the past 6 months. HIV-uninfected women were less likely than HIV-infected women to use condoms in the past 6 months (*P* = .001). Self-report of *Trichomonas vaginalis* infection, BV, candidiasis, or STIs within the past

**Table 1. Selected Characteristics of Participants in the Cervicovaginal Microbiota Substudy at their Index Visit, Stratified by Study Group**

Characteristic	HIV Uninfected (n = 22)	HIV Infected, Stable Immune Status (n = 22)	HIV Infected, Rapid Immunosuppression (n = 20)	P Value <sup>a</sup>	
				Between All Groups	Between HIV- Positive Groups
Age, y, mean $\pm$ SD	30.7 $\pm$ 9.3	33.7 $\pm$ 6.8	31.9 $\pm$ 4.9	.41	.35
CD4 <sup>+</sup> T-cell count, cells/ $\mu$ L, median (IQR)	1150 (902–1428)	782 (556–918)	285 (150–463)	<.001	<.001
No. of male sex partners in past 6 mo				.87	.45
0	4 (18)	3 (14)	5 (25)		
1	12 (55)	11 (50)	12 (60)		
2	4 (18)	5 (23)	2 (10)		
$\geq 3$	2 (9)	3 (14)	1 (5)		
Lifetime no. of male sex partners				.28	.34
<5	1 (5)	2 (10)	6 (30)		
5–9	6 (27)	2 (10)	3 (15)		
10–49	8 (36)	9 (43)	5 (25)		
$\geq 50$	7 (32)	8 (38)	6 (30)		
Male condom use in past 6 mo				.001	.45
No	14 (64)	3 (14)	5 (25)		
Yes	8 (36)	19 (86)	15 (75)		
Smoking				.15	.08
Never smoked	5 (23)	3 (14)	8 (40)		
Former smoker	3 (14)	1 (5)	0 (0)		
Current smoker	14 (64)	18 (82)	12 (60)		
Oral contraceptive use in past 6 mo				.68	.60
No	15 (83)	16 (94)	14 (82)		
Yes	3 (17)	1 (6)	3 (18)		
Self-reported <i>T. vaginalis</i> infection in past 6 mo				.56	.87
No	18 (86)	16 (73)	15 (75)		
Yes	3 (14)	6 (27)	5 (25)		
Self-reported BV in past 6 mo				.22	.11
No	20 (91)	21 (100)	17 (85)		
Yes	2 (9)	0 (0)	3 (15)		
Self-reported <i>Candida</i> infection in past 6 mo				.49	.76
No	13 (59)	10 (45)	8 (40)		
Yes	9 (41)	12 (55)	12 (60)		
Self-reported no. of STIs in past 6 mo				.65	.96
0	10 (45)	12 (55)	12 (60)		
1	6 (27)	2 (9)	2 (10)		
2	4 (18)	3 (14)	3 (15)		
$\geq 3$	2 (9)	5 (23)	3 (15)		
Any HPV infection				.02	.06
No	17 (77)	14 (64)	7 (35)		
Yes	5 (23)	8 (36)	13 (65)		
Oncogenic HPV infection				.04	.22
No	21 (95)	18 (82)	13 (65)		
Yes	1 (5)	4 (18)	7 (35)		
Vaginal pH, median (IQR)	5.0 (4.7–5.5)	5.0 (4.4–5.5)	5.0 (4.0–6.0)	.79	.52

Data are no. (%) of participants or median value (interquartile range). See “Methods” section for details (substudy population). Briefly, 3 study groups were defined: HIV-uninfected women (n = 22); HIV-infected women with stable immune status, defined as a CD4<sup>+</sup> T-cell count of  $>500$  cells/mm<sup>3</sup> throughout the observation period, individually matched at the index visit to HIV-infected women with rapidly progressive immunosuppression (n = 22); and HIV-infected women with rapidly progressive immunosuppression, defined as an initial CD4<sup>+</sup> T-cell count of  $>500$  cells/mm<sup>3</sup> followed by a decrease to  $<200$  cells/mm<sup>3</sup>, with a poor immunologic response to therapy ( $<100$  cells/mm<sup>3</sup> increase over a year) during the observation period (n = 20). However, as the case definition in the current substudy is related to HPV and not HIV risk group, the matched-pairs were no longer relevant, and we instead controlled for these factors as covariates in our models. Abbreviations: BV, bacterial vaginosis; HIV, human immunodeficiency virus; HPV, human papillomavirus; IQR, interquartile range; STI, sexually transmitted infection; *T. vaginalis*, *Trichomonas vaginalis*.

<sup>a</sup> By analysis of variance (for continuous variables with normal distribution), the Kruskal–Wallis test (for continuous variables with nonnormal distribution), or the  $\chi^2$  test (for categorical data).



**Table 2. Spearman Correlations Between Cervicovaginal pH and the Relative Abundance of Lactobacilli, by Genus and Species**

<i>Lactobacillus</i>	Spearman Correlation	P Value
Genus	−0.67	<.001
<i>L. crispatus</i>	−0.48	<.001
<i>L. gasseri</i>	−0.39	<.001
<i>L. iners</i>	−0.32	<.001
<i>L. jensenii</i>	−0.40	<.001

6 months did not differ between the 3 groups, although by design subjects were excluded if laboratory testing at the index visit detected gonorrhea, syphilis, trichomoniasis, or chlamydial infection. The median number of CVLs tested for both microbiome and HPV in the current study was 6 (range, 2–11); the median follow-up time was 82.8 person-months (range, 15.6–137.7 person-months).

Table 2 shows Spearman correlations between *Lactobacillus* species RA and pH. The RA of each *Lactobacillus* species was inversely associated with pH. The correlation coefficients for any individual species of *Lactobacillus* was greatest for *L. crispatus* ( $r = -0.48$ ), although the highest correlation we observed was for *Lactobacillus* species (ie, at the genus level;  $r = -0.67$ ).

Hierarchical clustering found 6 CSTs, with 5 dominated by an individual *Lactobacillus* species, whereas CST IV-B comprised a variety of genera (eg, *Prevotella*, *Atopobium*, *Megasphaera*, and *Sneathia*) with a low RA of *Lactobacillus* species. Figure 2 displays a heat map of the RA data for the 25 most abundant genera or species and how they cluster within each CST. *L. crispatus* dominated CST I, CST II was dominated by *Lactobacillus gasseri*, both CST III and IV-A were dominated by *Lactobacillus iners*, and CST V was dominated by *Lactobacillus jensenii*. Figure 2 additionally shows the SDI for each study visit, and as expected CSTs I through IV-A and V (each dominated by a single *Lactobacillus* species) had a low SDI, compared with CST IV-B. The figure also presents the proportion of oncogenic HPV-positive women-visits for each CST.

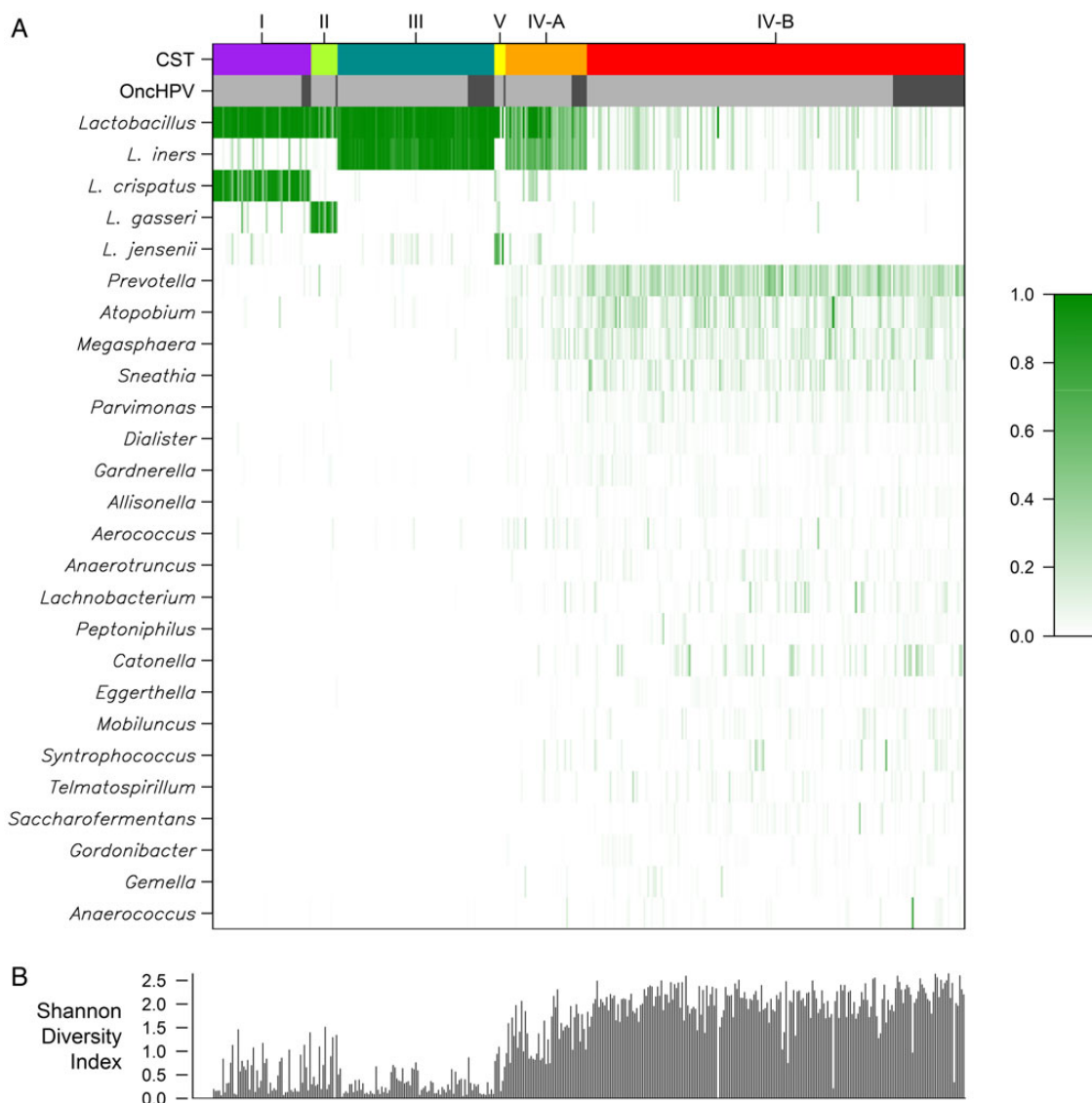
Table 3 shows multivariate GEE logistic regression models of the associations between CSTs and the prevalent detection of HPV. These models adjusted for HIV study group, time-updated covariates, and intra-individual correlations between repeated cross-sectional measures of association with prevalent HPV (see “Methods” section). While CST I (*L. crispatus* dominated) had an inverse association with detection of any HPV and especially with any oncogenic HPV, the relationship was of borderline statistical significance (eg, odds ratio [OR] for oncogenic HPV was 0.30 [95% confidence interval {CI}, .08–1.11];  $P = .07$ ). In contrast, when we assessed the RA of individual *Lactobacillus* species, characterized in an ordinal fashion (Table 4), we observed stronger, statistically significant associations with detection of HPV. Briefly, high (vs low) RA of *L. crispatus* was significantly associated with a lower relative odds of any

HPV (OR, 0.48; 95% CI, .24–.96;  $P_{\text{trend}} = .03$ ) or any oncogenic HPV (OR, 0.14; 95% CI, .01–1.59;  $P_{\text{trend}} = .03$ ). However, no associations were found between any HPV or oncogenic HPV detection and the RA of *L. iners*, *L. gasseri*, or *L. jensenii* (data not shown). In addition, no association between pH and HPV detection was observed in any of the above multivariate models, regardless of how pH was parameterized, including as a continuous or as an ordinal based on Amsel criteria. Nor did exclusion of pH from the above models discernibly alter any findings (data not shown). When we assessed SDI tertile instead of CST or *Lactobacillus* species in multivariate GEE logistic regression models, we found significantly higher oncogenic HPV detection in women in the highest SDI tertile (OR, 2.04; 95% CI, 1.18–3.52;  $P_{\text{trend}} = .01$ ), relative to those in the lowest SDI tertile.

Last, in an exploratory analysis, we used continuous-time multistate Markov models fit using maximum likelihood methods to assess whether transition to a CST dominated by *L. crispatus* was associated with a change in the risk of incident detection of HPV (ie, HPV type[s] not found in any earlier CVLs from the same woman). The transition rate ratio (TRR), a value that can be thought of as similar to a hazard ratio [15], was 0.17 (95% CI, .04–.79;  $P = .02$ ) after adjustment for study group, and this effect estimate was little changed (TRR, 0.20; 95% CI, .03–1.14;  $P = .07$ ) by adjustment for additional covariates, indicating that the most appropriate model was the parsimonious model adjusted only for study group.

## DISCUSSION

In the current study, high *L. crispatus* RA was associated with decreased prevalence of oncogenic HPV in HIV-infected and HIV-uninfected women across multiple serial annual visits, even after statistical adjustment for multiple cofactors, including host immune status (as defined by HIV study group) and vaginal pH. We predicted these results, because earlier (largely cross-sectional) studies had found inverse associations between *L. crispatus* abundance and the risks of several STIs [32–34]. Further, the one previous longitudinal study of cervicovaginal microbiota and HPV, as well as cross-sectional studies in Rwanda and South Africa, each reported inverse associations of HPV prevalence with CSTs dominated by *L. crispatus* or with *L. crispatus* RA [15, 32]. To our knowledge, the current study is the first investigation involving serially repeated molecular characterization of the vaginal microbiome and its relation with oncogenic HPV in HIV-infected women. This allowed us to conduct an exploratory analysis in which we prospectively assessed transitions between CSTs over time and found that transition to an *L. crispatus*-dominated CST was associated with a statistically significant reduction in the risk of a new (previously undetected) HPV type. If correct, these data could reflect either a reduction in the rate of newly acquired (incident) HPV infection or reduced recurrence/reactivation of previously acquired HPV.



**Figure 2.** Heat map of microbiome relative abundances and Shannon diversity index. *A*, Heat map of the relative percent abundance of bacterial taxa, by community state type (CST) and the presence of oncogenic human papillomavirus (oncHPV). The results reflect the 25 most abundant bacterial taxa. Each vertical line represents a person-visit, with relative abundance indicated by the shade of green as shown in the key. Results for CST (multicolor) and oncHPV (dark gray, present; light gray, absent) are shown across the top. The CST categories (CST I–V) used are similar to those in prior reports and ordered in the figure in a similar manner [15]. *B*, Shannon diversity index calculated for each of the 398 observations.

These findings have potential clinical implications, since they raise the possibility that promoting high cervicovaginal levels of *L. crispatus* (eg, via a vaginal suppository) may help reduce the burden of oncogenic HPV in women. The implications may be especially important for HIV-infected women, since the inverse *L. crispatus*–oncogenic HPV associations we observed were independent of HIV and host immune status and therefore indicate similar potential benefit among those with a low CD4<sup>+</sup> T-cell count.

We additionally found that a low SDI was associated with reduced prevalent detection of oncogenic HPV. While this is consistent with an inverse association of oncogenic HPV with

an *L. crispatus*–dominated CST (which had a low SDI), it could also reflect an inverse association of oncogenic HPV detection with other low SDI CSTs or a positive association with a high SDI CST, such as CST IV-B. However, additional analyses failed to show significant associations of oncogenic HPV (or any HPV) with CSTs or *Lactobacillus* species beyond the *L. crispatus*–oncogenic HPV relationship.

The current results were also noteworthy for the lack of association of vaginal pH with oncogenic HPV detection in multivariate models. While there has been considerable discussion in the literature regarding the potential antiviral effects of low pH and its possible role in explaining the beneficial activity of

**Table 3. Relation of Vaginal Community State Types (CSTs) With Detection of Human Papillomavirus (HPV)**

Variable	Model 1 (pH Continuous)		Model 2 (pH Binary Variable)		Model 3 (pH Ordinal With 3 Levels)	
	OR (95% CI)	P Value	OR (95% CI)	P Value	OR (95% CI)	P Value
<b>Any HPV</b>						
IV-B, low <sup>a</sup> <i>Lactobacillus</i>	Reference		Reference		Reference	
I, <i>L. crispatus</i>	0.58 (.29–1.16)	.12	0.59 (.30–1.19)	.14	0.57 (.28–1.15)	.12
II, <i>L. gasseri</i>	0.84 (.44–1.61)	.60	0.82 (.43–1.55)	.54	0.80 (.42–1.51)	.49
III, <i>L. iners</i>	1.10 (.72–1.69)	.66	1.09 (.71–1.69)	.68	1.04 (.68–1.60)	.84
IV-A, medium <sup>a</sup> <i>L. iners</i>	1.34 (.82–2.19)	.24	1.27 (.80–2.02)	.31	1.22 (.75–1.99)	.43
V, <i>L. jensenii</i>	1.12 (.61–2.08)	.71	1.02 (.60–1.72)	.95	0.96 (.53–1.73)	.89
Vaginal pH <sup>b</sup>						
Continuous	1.19 (.95–1.47)	.12	. . .		. . .	
>4.5 (vs ≤4.5)	. . .		1.40 (.94–2.07)	.09	. . .	
Medium (vs low)	. . .		. . .		1.44 (.99–2.09)	.06
High (vs low)	. . .		. . .		1.26 (.73–2.18)	.40
<b>Oncogenic HPV</b>						
IV-B, low <sup>a</sup> <i>Lactobacillus</i>	Reference		Reference		Reference	
I, <i>L. crispatus</i>	0.34 (.09–1.33)	.12	0.29 (.08–1.05)	.06	0.30 (.08–1.11)	.07
II, <i>L. gasseri</i>	0.61 (.06–6.00)	.67	0.55 (.06–5.13)	.60	0.57 (.06–5.32)	.62
III, <i>L. iners</i>	1.42 (.66–3.04)	.37	1.19 (.69–2.04)	.53	1.26 (.66–2.38)	.49
IV-A, medium <sup>a</sup> <i>L. iners</i>	1.84 (.78–4.34)	.16	1.65 (.72–3.78)	.24	1.73 (.70–4.29)	.23
V, <i>L. jensenii</i>	1.09 (.37–3.18)	.88	0.84 (.33–2.17)	.72	0.90 (.33–2.50)	.84
Vaginal pH <sup>b</sup>						
Continuous	1.28 (.85–1.93)	.24	. . .		. . .	
>4.5 (vs ≤4.5)	. . .		1.26 (.67–2.37)	0.46	. . .	
Medium (vs low)	. . .		. . .		1.21 (.66–2.23)	.53
High (vs low)	. . .		. . .		1.40 (.56–3.54)	.47

All models were adjusted for age, study group, highly active antiretroviral therapy use, number of recent sex partners, smoking status, condom use, lifetime sex partners, *Trichomonas vaginalis* infection in the past 6 months, *Candida* infection in the past 6 months, or other sexually transmitted infections (STIs) in the past 6 months. Those with an STI on the day of the index visit were excluded from this substudy.

Abbreviations: BV, bacterial vaginosis; CI, confidence interval; OR, odds ratio.

<sup>a</sup>“Low” and “medium” denote the relative abundance of *Lactobacillus* species and *L. iners*, respectively.

<sup>b</sup> A pH of ≤4.5 (the cut point used in the Amsel criteria for the diagnosis of BV) was considered low, a pH of 4.5–5.8 was considered medium, and a pH of >5.8 (where a pH of 5.8 was the median of all pH values >4.5) was considered high.

*Lactobacillus* species against STIs, most of that research preceded widespread use of methods for comprehensively assessing the cervicovaginal microbiome [11]. The growing evidence of species-specific (eg, *L. crispatus*–related) relationships between the cervicovaginal microbiome and HPV infection argues against a significant role of pH, since production of lactic acid is a common attribute of lactobacilli as a group, and our data showed that the greatest association with pH was with the RA of *Lactobacillus* species, not for *L. crispatus* RA or CSTs dominated by *L. crispatus*. Thus, biologic mechanisms other than pH are likely involved in the anti-HPV activity of *L. crispatus*.

For example, recent data suggest that the 2 isoforms of lactic acid, D-lactic acid and L-lactic acid, might differ in their microbial effects. D-lactic acid (but not L-lactic acid) levels have been inversely associated with the ability of HIV to transverse cervicovaginal mucous [35]. An antiviral effect of D-lactic acid would be consistent with our results for oncogenic HPV, since *L. crispatus* produces D and L isoforms of lactic acid, whereas *L. iners*

produces only L-lactic acid [35, 36]. *L. gasseri* and *L. jensenii* also produce D-lactic acid, and their RAs were nonsignificantly inversely associated with HPV detection in the current study. Thus, high D-lactic acid levels but not acidity per se could be associated with reduced infection and/or transmission of certain viruses, including HPV.

While production of H<sub>2</sub>O<sub>2</sub> has been observed in the majority of *L. crispatus* isolates, prior data from our group suggest that the levels of H<sub>2</sub>O<sub>2</sub> produced are likely too limited to greatly affect the risk of STIs [7], especially HPV infection, since HPV is a nonenveloped virus and, therefore, likely to be resistant to H<sub>2</sub>O<sub>2</sub>. On the other hand, bacteriocins produced by *L. crispatus* may have antiviral effects. Bacteriocins are ribosomally synthesized antimicrobial peptides and proteins [37, 38]. Relatively little has been written regarding the bacteriocins produced by cervicovaginal *Lactobacillus* species, with most of the bacteriocin literature to date focused on *Lactobacillus* species inhabiting the gut. *L. crispatus* has additionally been noted to have potential immunomodulatory activity. For

**Table 4. Relation of *Lactobacillus crispatus* Relative Abundance to Detection of Human Papillomavirus (HPV)**

Characteristic	Model 1 (pH Continuous)		Model 2 (pH Binary Variable)		Model 3 (pH Ordinal With 3 Levels)	
	OR (95% CI)	P Value	OR (95% CI)	P Value	OR (95% CI)	P Value
<b>Any HPV</b>						
<i>L. crispatus</i> relative abundance						
Low	Reference	.02 <sup>a</sup>	Reference	.03 <sup>a</sup>	Reference	.03 <sup>a</sup>
Medium	0.52 (.24–1.11)	.09	0.54 (.26–1.14)	.11	0.53 (.25–1.12)	.10
High	0.47 (.24–.90)	.02	0.49 (.25–.96)	.04	0.48 (.24–.96)	.04
Vaginal pH <sup>b</sup>						
Continuous	1.11 (.91–1.35)	.32	. . .		. . .	
>4.5 (vs ≤4.5)	. . .		1.28 (.86–1.90)	.23	. . .	
Medium (vs low)	. . .		. . .		1.37 (.95–1.98)	.09
High (vs low)	. . .		. . .		1.15 (.70–1.91)	.58
<b>Oncogenic HPV</b>						
<i>L. crispatus</i> relative abundance						
Low	Reference	.04 <sup>a</sup>	Reference	.03 <sup>a</sup>	Reference	.03 <sup>a</sup>
Medium	0.24 (.07–.81)	.02	0.23 (.07–.74)	.01	0.23 (.07–.74)	.01
High	0.14 (.01–1.65)	.12	0.14 (.01–1.58)	.11	0.14 (.01–1.59)	.11
Vaginal pH <sup>b</sup>						
Continuous	1.09 (.78–1.53)	.63	. . .		. . .	
>4.5 (vs ≤4.5)	. . .		1.11 (.56–2.18)	.77	. . .	
Medium (vs low)	. . .		. . .		1.11 (.54–2.29)	.77
High (vs low)	. . .		. . .		1.10 (.51–2.41)	.81

All models were adjusted for age, study group, highly active antiretroviral therapy use, number of recent sex partners, smoking status, condom use, lifetime sex partners, *Trichomonas vaginalis* infection in the past 6 months, *Candida* infection in the past 6 months, or other sexually transmitted infections (STIs) in the past 6 months. Those with an STI on the day of the index visit were excluded from this substudy.

Abbreviations: CI, confidence interval; OR, odds ratio.

<sup>a</sup> P for trend.

<sup>b</sup> A pH of ≤4.5 (the cut point used in the Amsel criteria for the diagnosis of BV) was considered low, a pH of 4.5–5.8 was considered medium, and a pH of >5.8 (where a pH of 5.8 was the median of all pH values >4.5) was considered high.

example, *Candida albicans* growth in vitro was inhibited via modulation of Toll-like receptors 2/4 by *L. crispatus* in epithelial cell culture [39].

While the current study contributes data suggesting a possible protective effect of *L. crispatus* against cervicovaginal HPV infection in HIV-infected and HIV-uninfected women (independent of pH), several limitations need to be considered. These include the relatively small number of HIV-infected and HIV-uninfected subjects, even though longitudinal testing of nearly 400 CVLs was performed. This was addressed in part by the use of efficient analytic methods able to incorporate all available serial HPV and microbiome data. However, larger prospective studies will be needed to more comprehensively study the relation of the cervicovaginal microbiome to the full natural history of HPV, including the presence and development of cervical disease. These clinical/pathologic end points were too infrequent in the current data set to adequately study. Nor did we have sufficient data to assess statistical interaction (eg, by immune strata), and it may require subanalyses among sexually inactive women to help differentiate reactivation of previously acquired from newly acquired (incident) HPV. The study also focused on only premenopausal African American women, which was done to reduce the number of factors being studied.

Thus, future studies will need to include women with a broader range of age, menopausal status, and race/ethnicity.

In summary, our results suggest a beneficial effect of *L. crispatus* on the burden of HPV in HIV-infected and HIV-uninfected women that is unrelated to pH. Further research to better understand these relationships is strongly warranted, including large prospective cohort studies and laboratory research to understand the mechanisms that underlie the relation of the cervicovaginal microbiota with HPV infection. However, if the current data are confirmed in large prospective cohort studies, randomized clinical trials of *L. crispatus* (eg, as a probiotic vaginal suppository) to reduce HPV and cervical disease might be appropriate, even if the mechanisms remain uncertain. The clinical implications could be especially important for HIV-infected women, as the oncogenic HPV–*L. crispatus* relationship appeared to be independent of host immune status.

## Notes

**Financial support.** This work was supported by the National Cancer Institute (NCI; grants R01-CA-085178 and R01-CA-174634 to H. D. S. and grant P30-CA-013330); C. P. Z. was supported by NIH 5K12GM102779; the American Recovery Reinvestment Act, National Institute of Allergy and Infectious Diseases (NIAID) supplement to the Chicago Women's Interagency HIV Study (WIHS; to G. S.); the NIAID (grant



P30-AI-051519); and the WIHS. WIHS (principal investigators): UAB-MS WIHS (Michael Saag, Mirjam-Colette Kempf, and Deborah Konkle-Parker), U01-AI-103401; Atlanta WIHS (Ighowherha Ofotokun and Gina Wingood), U01-AI-103408; Bronx WIHS (Kathryn Anastos), U01-AI-035004; Brooklyn WIHS (Howard Minkoff and Deborah Gustafson), U01-AI-031834; Chicago WIHS (Mardge Cohen), U01-AI-034993; Metropolitan Washington WIHS (Mary Young), U01-AI-034994; Miami WIHS (Margaret Fischl and Lisa Metsch), U01-AI-103397; UNC WIHS (Adaora Adimora), U01-AI-103390; Connie Wofsy Women's HIV Study, Northern California (Ruth Greenblatt, Bradley Aouizerat, and Phyllis Tien), U01-AI-034989; WIHS Data Management and Analysis Center (Stephen Gange and Elizabeth Golub), U01-AI-042590; Southern California WIHS (Alexandra Levine and Marek Nowicki), U01-HD-032632 (WIHS I-WIHS IV). The WIHS is funded primarily by the NIAID, with additional co-funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, the NCI, the National Institute on Drug Abuse, and the National Institute of Mental Health. Targeted supplemental funding for specific projects is also provided by the National Institute of Dental and Craniofacial Research, the National Institute on Alcohol Abuse and Alcoholism, the National Institute on Deafness and other Communication Disorders, and the National Institutes of Health Office of Research on Women's Health. WIHS data collection is also supported by UL1-TR000004 (UCSF CTSA) and UL1-TR000454 (Atlanta CTSA).

**Potential conflicts of interest.** H. D. S. reports leading a study involving free blinded testing using HPV E6/E7 protein assays by Arbor Vita, p16/Ki67 cytology by MTM Laboratories/Ventura-Roche, and MCM-2/TOP2A cytology by BD Diagnostics; no financial payments to H. D. S. or his home institution were received. K. A. has received honoraria from Bristol-Myers Squibb. J. M. P. reports grants and nonfinancial support from Merck and Hologic, reports nonfinancial support from Vaxgenetics, and has been compensated financially by Bristol-Myers Squibb. All other authors report no potential conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

## References

- Wallin KL, Wiklund F, Angstrom T, et al. Type-specific persistence of human papillomavirus DNA before the development of invasive cervical cancer. *N Engl J Med* **1999**; 341:1633–8.
- Abraham AG, D'Souza G, Jing Y, et al. Invasive cervical cancer risk among HIV-infected women: a North American multicohort collaboration prospective study. *J Acquir Immune Defic Syndr* **2013**; 62:405–13.
- Palefsky JM, Minkoff H, Kalish LA, et al. Cervicovaginal human papillomavirus infection in human immunodeficiency virus-1 (HIV)-positive and high-risk HIV-negative women. *J Natl Cancer Inst* **1999**; 91:226–36.
- Massad LS, Xie X, Burk R, et al. Long-term cumulative detection of human papillomavirus among HIV seropositive women. *AIDS* **2014**; 28:2601–8.
- Strickler HD, Burk RD, Fazzari M, et al. Natural history and possible reactivation of human papillomavirus in human immunodeficiency virus-positive women. *J Natl Cancer Inst* **2005**; 97:577–86.
- Watts DH, Fazzari M, Minkoff H, et al. Effects of bacterial vaginosis and other genital infections on the natural history of human papillomavirus infection in HIV-1-infected and high-risk HIV-1-uninfected women. *J Infect Dis* **2005**; 191:1129–39.
- Mirmonsef P, Hotton AL, Gilbert D, et al. Free glycogen in vaginal fluids is associated with *Lactobacillus* colonization and low vaginal pH. *PLoS One* **2014**; 9:e102467.
- Ravel J, Gajer P, Abdo Z, et al. Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A* **2011**; 108(suppl 1):4680–7.
- Clarke MA, Rodriguez AC, Gage JC, et al. A large, population-based study of age-related associations between vaginal pH and human papillomavirus infection. *BMC Infect Dis* **2012**; 12:33.
- Mehta SD, Donovan B, Weber KM, et al. The vaginal microbiota over an 8- to 10-year period in a cohort of HIV-infected and HIV-uninfected women. *PLoS One* **2015**; 10:e0116894.
- Nardis C, Mosca L, Mastromarino P. Vaginal microbiota and viral sexually transmitted diseases. *Ann Ig* **2013**; 25:443–56.
- Benning L, Golub ET, Anastos K, et al. Comparison of lower genital tract microbiota in HIV-infected and uninfected women from Rwanda and the US. *PLoS One* **2014**; 9:e96844.
- Cohen CR, Lingappa JR, Baeten JM, et al. Bacterial vaginosis associated with increased risk of female-to-male HIV-1 transmission: a prospective cohort analysis among African couples. *PLoS Med* **2012**; 9:e1001251.
- Jespersen V, Menten J, Smet H, et al. Quantification of bacterial species of the vaginal microbiome in different groups of women, using nucleic acid amplification tests. *BMC Microbiol* **2012**; 12:83.
- Brotman RM, Shardell MD, Gajer P, et al. Interplay between the temporal dynamics of the vaginal microbiota and human papillomavirus detection. *J Infect Dis* **2014**; 210:1723–33.
- Barkan SE, Melnick SL, Preston-Martin S, et al. The Women's Interagency HIV Study. WIHS Collaborative Study Group. *Epidemiology* **1998**; 9:117–25.
- Ho DE, Imai K, King G, Stuart EA. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Polit Analysis* **2007**; 15:199–236.
- Ho DE, Imai K, King G, Stuart EA. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *J Stat Softw* **2011**; 42:1–28.
- Burk RD, Ho GY, Beardsley L, Lempa M, Peters M, Bierman R. Sexual behavior and partner characteristics are the predominant risk factors for genital human papillomavirus infection in young women. *J Infect Dis* **1996**; 174:679–89.
- Goldberg GL, Vermund SH, Schiffman MH, Ritter DB, Spitzer C, Burk RD. Comparison of Cytobrush and cervicovaginal lavage sampling methods for the detection of genital human papillomavirus. *Am J Obstet Gynecol* **1989**; 161:1669–72.
- Gravitt PE, van Doorn LJ, Quint W, et al. Human papillomavirus (HPV) genotyping using paired exfoliated cervicovaginal cells and paraffin-embedded tissues to highlight difficulties in attributing HPV types to specific lesions. *J Clin Microbiol* **2007**; 45:3245–50.
- IARC. IARC monographs on the evaluation of carcinogenic risks to humans: human papillomaviruses. Vol 90. Lyon, France: World Health Organization, **2007**.
- Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **2010**; 7:335–6.
- Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **2007**; 73:5261–7.
- Cole JR, Chai B, Farris RJ, et al. The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. *Nucleic Acids Res* **2005**; 33:D294–6.
- Milligan GW. An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika* **1980**; 45:325–42.
- Milligan GW. A Monte Carlo study of thirty internal criterion measures for cluster analysis. *Psychometrika* **1981**; 46:187–99.
- Hill MO. Diversity and evenness: a unifying notation and its consequences. *Ecology* **1973**; 54:427–32.
- Xue X, Gange SJ, Zhong Y, et al. Marginal and mixed-effects models in the analysis of human papillomavirus natural history data. *Cancer Epidemiol Biomarkers Prev* **2010**; 19:159–69.
- Moore SH, Boot HJ, Speksnijder AG, et al. Six-month incidence and persistence of oral HPV infection in HIV-negative and HIV-infected men who have sex with men. *PLoS One* **2014**; 9:e98955.
- Louvanto K, Rautava J, Syrjanen K, Grenman S, Syrjanen S. The clearance of oral high-risk human papillomavirus infection is impaired by long-term persistence of cervical human papillomavirus infection. *Clin Microbiol Infect* **2014**; 20:1167–72.
- Borgdorff H, Tsvitvadze E, Verhelst R, et al. *Lactobacillus*-dominated cervicovaginal microbiota associated with reduced HIV/STI prevalence and genital HIV viral load in African women. *ISME J* **2014**; 8:1781–93.
- Petrova MI, van den Broek M, Balzarini J, Vanderleyden J, Lebeer S. Vaginal microbiota and its role in HIV transmission and infection. *FEMS Microbiol Rev* **2013**; 37:762–92.
- van de Wijgert JH, Borgdorff H, Verhelst R, et al. The vaginal microbiota: what have we learned after a decade of molecular characterization? *PLoS One* **2014**; 9:e105998.
- Nunn KL, Wang YY, Harit D, et al. Enhanced Trapping of HIV-1 by Human Cervicovaginal Mucus Is Associated with *Lactobacillus crispatus*-Dominant Microbiota. *MBio* **2015**; 6:e01084–15.
- Witkin SS. The vaginal microbiome, vaginal anti-microbial defence mechanisms and the clinical challenge of reducing infection-related preterm birth. *BJOG* **2015**; 122:213–8.
- Cotter PD, Ross RP, Hill C. Bacteriocins—a viable alternative to antibiotics? *Nat Rev Microbiol* **2013**; 11:95–105.
- Petrova MI, Lievens E, Malik S, Imholz N, Lebeer S. *Lactobacillus* species as biomarkers and agents that can promote various aspects of vaginal health. *Front Physiol* **2015**; 6:81.
- Rizzo A, Losacco A, Carratelli CR. *Lactobacillus crispatus* modulates epithelial cell defense against *Candida albicans* through Toll-like receptors 2 and 4, interleukin 8 and human beta-defensins 2 and 3. *Immunol Lett* **2013**; 156:102–9.