Sex work is associated with increased vaginal microbiome diversity in young women from Mombasa, Kenya

Aida Sivro, PhD,1,2* Ruth Mwatelah, MSc,2* Cheli Kambaran, BSc,2* Henok Gebrebrhan, BSc,2 Michael G. Becker, PhD,3, Huiting Ma, PhD,4,5 Nichole R. Klatt, PhD,6,7 Alexander S. Zevin, PhD,7 Nzioki King’ola, MA,8 Sammy Wambua, MSc,8 Peter Gichangi, MBChB MMed,8 Eve Cheuk, PhD,9 Paul J. McLaren, PhD,2,3 Sharmistha Mishra, MD PhD,3,4 Marissa Becker, MD MSc,2,9 Lyle R. McKinnon, PhD,1,2,3,10

* Equal contributors

1 Centre for the AIDS Programme of Research in South Africa (CAPRISA), Durban, South Africa
2 Department of Medical Microbiology and Infectious Diseases, University of Manitoba, Winnipeg, Canada
3 JC Wilt Infectious Disease Research Centre, Public Health Agency of Canada, Winnipeg, Canada.
4 Department of Medicine, University of Toronto, Toronto, Canada.
5 MAP Centre for Urban Health Solutions, Li Ka Shing Knowledge Institute, University of Toronto, Toronto, Canada
6 Department of Pediatrics, Miller School of Medicine, University of Miami, Miami, FL, USA
7 Department of Pharmaceutics, University of Washington, Seattle, WA, USA
8 International Centre for reproductive Health, Mombasa, Kenya.
9 Centre for Global Public Health, University of Manitoba, Winnipeg, Canada.
10 Department of Medical Microbiology, University of Nairobi, Nairobi, Kenya.
Corresponding author: Lyle R. McKinnon, Department of Medical Microbiology and Infectious Diseases, University of Manitoba, Room 504-745 Bannatyne Ave, Winnipeg, MB CANADA R3E 0J9. Phone: (204) 975-7708. Email: lyle.mckinnon@umanitoba.ca

Running head: Sex work & vaginal microbiome diversity

Word count, main text: 3,285

Word count, abstract: 250

Conflict of interests: The authors declare no conflicts of interest.

Prior presentation. This work was presented in part at the Annual Meeting of the Canadian Association of HIV Researchers (CAHR) in Vancouver, Canada, April 2018.

Funding statement: The Transitions study was funded by an operating grant (MOP-13044) from the Canadian Institutes of Health Research (CIHR). Biological aspects of the study were funded by the Canadian Institute of Health Research (CIHR), grant numbers TMI 138658 and PJT-148796. AS is funded by the European & Developing Countries Clinical Trials Partnership (EDCTP) Career Development Fellowship (TMA2016CDF-1582); SM is supported by a CIHR-OHTN New Investigator Award; MB and LRM are supported by CIHR New Investigator Awards.

Acknowledgments: We would like to thank the study participants, International Centre for Reproductive Health management (Mary Thion’o), clinic (Agnes Rinyiru, Bridgit Mrima, and Jacqueline Chokwe) and laboratory staff that participated in the Transitions study in Mombasa, Kenya. We also thank the team at Partners in Health and Development in Africa (PHDA) in Nairobi, who assisted with data entry, and the team at the Public Health Agency of Canada (PHAC) who provided serological testing data.
Author contributions:

Designed the study: RM, EC, SM, MB, LRM

Supervised field work: RM, NK, SW, PG

Carried out cytokine analyses: AS

Microbiome data acquisition: CK, NRK, ASZ

Analysed the data: AS, RM, CK, HG, MGB, HM, PJM, LRM

Wrote and reviewed the paper: AS, RM, CK, HG, MGB, HM, NRK, ASZ, NK, SW, PG, EC, PJM, SM, MB, LRM

Data Availability Statement: All data will be made available upon request to the corresponding author.

Abstract

Background. While non-optimal vaginal bacteria and inflammation have been associated with increased HIV risk, the upstream drivers of these phenotypes are poorly defined in young African women.

Setting. Mombasa, Kenya.

Methods. We characterized vaginal microbiome and cytokine profiles of sexually active young women aged 14-24 years (n=168) in three study groups: those engaging in formal sex work, in transactional sex, and non-sex workers. Vaginal secretions were collected via self-inserted SoftCup™, and assayed for cytokines and vaginal microbiome via multiplex ELISA and 16S rRNA sequencing, respectively. Epidemiological data were captured using a validated questionnaire.

Results. The median age of participants was 20 (IQR: 18-22). Approximately two-thirds of young women (105/168) had vaginal microbial communities characterized by Gardnerella and/or Prevotella spp.-dominance; a further 29% (49/168) were predominantly Lactobacillus iners. Microbiome clustering explained a large proportion of cytokine variation (>50% by the first 2
principle components). Age was not associated with vaginal microbial profiles in bivariable or multivariable analyses. Women self-identifying as sex workers had increased alpha (intra-individual) diversity, independent of age, recent sexual activity, HIV and other STIs (beta = 0.47, 95% CI: 0.05 – 0.90, p = 0.03). Recent sex (number of partners or sex acts last week, time since last vaginal sex) correlated with increased alpha diversity, particularly in participants who were not involved in sex work.

**Conclusion.** Non-optimal vaginal microbiomes were common in young Kenyan women and associated with sex work and recent sexual activity, but independent of age. Restoring optimal vaginal microflora may represent a useful HIV prevention strategy.

**Key words.** HIV; female sex worker; inflammation; vaginal microbiome; bacterial vaginosis; young women

**Introduction**

HIV is often transmitted across genital mucosa, where the virus must infect susceptible HIV target cells\(^1\). Before HIV can access these cells, the virus must first survive mucosal defenses and cross the robust mucosal barrier that prevents HIV transmission during most exposures\(^2\). Genital inflammation defined by elevated inflammatory cytokines in genital secretions has been associated with HIV acquisition risk\(^3\) and reduced efficacy of vaginally applied tenofovir\(^4\). However, the causes of genital inflammation remain poorly defined. Contributing upstream drivers likely include variables that have been associated with both increased HIV risk and elevated cytokines; these include sexually transmitted infections (STIs), abrasion, contraceptives, semen exposure, dysbiosis and vaginal douching\(^5,6\).

In addition to the above, recent studies suggest that genital inflammation is strongly associated with bacterial communities dominated by anaerobes, such as *Gardnerella* and *Prevotella* spp., amongst many others\(^7-9\). These organisms thrive in the absence of healthy *Lactobacilli* (non-iners) spp.,
which if present help to maintain lactic acid production and a low vaginal pH. Microbiologically, this is often defined as bacterial vaginosis (BV), a condition that overlaps with molecularly-defined BV, and has been associated with increased risk of HIV acquisition\textsuperscript{10,11}. A recent meta-analysis estimated that Nugent-defined BV affects almost one-third of women in Sub-Saharan Africa\textsuperscript{12}, including asymptomatic presentation in up to 85% of cases\textsuperscript{13}.

Despite the well-defined association between BV and HIV risk, there are limited data on the influence of sex work and sexual activity on the composition of vaginal microbiota, and resulting inflammation. One study of 32 adult women from Baltimore showed that sexual activity had a moderate influence on the composition of vaginal microbiota, independent of menstrual cycle phase\textsuperscript{14}. Multivariable analyses have shown that condom use, early sexual debut, age and history of STIs are correlated with molecular-BV (non-\textit{Lactobacillus} dominance as defined by bacterial sequencing)\textsuperscript{15}. However, few studies have focused on young women engaged in sex work, despite their increased risk of HIV acquisition\textsuperscript{16,17}. More studies are required in female sex workers (FSW) in order to capture the full spectrum of sexual activity, in the context of other variables, and its association with vaginal microbiome and inflammation profiles. In addition, most mucosal studies have focused on women \(\geq 18\) years of age, and therefore female genital tract immunology in younger women is poorly understood. The main objective of this study was to analyse the relationship between vaginal microbial community composition and mucosal immune milieu in young women in Kenya, including a subset <18 years of age, with varying patterns of sexual activity including a sub-group who self-identified as FSWs\textsuperscript{18}.

**Methods**

**Participant recruitment**

Participants for the parent Transitions study were recruited from hotspots, defined as locations where FSWs congregate to solicit clients, using multi-stage cluster sampling. Mapping of hotspots
was conducted in 2014 in Mombasa\textsuperscript{19}. From the total sample (n=1,299), we obtained specimens from 802 consenting participants; for this analysis we randomly selected 168 of these 802 participants with available mucosal specimens for cytokine and microbiome analyses. Peer educators recruited adolescent girls and young women 14-24 years of age via referral from hotspots to nearby study sites where trained research assistants conducted written informed consent and administered the questionnaire in local language. A series of questions triaged participants into three study categories; 1) self-identified sex workers, including those who had ever negotiated for sex prior to the act, 2) those engaging in transactional sex (had ever exchanged gifts for sex without prior negotiations) but who did not self-identify as sex workers and 3) a non-FSW group without participation in paid or transactional sex\textsuperscript{18}. The study was approved by the ethical review boards of the University of Manitoba and Kenyatta National Hospital.

**Sample collection and processing**

Mucosal samples were collected by participants by inserting a SoftCup\textsuperscript{TM} into the vagina for approximately 1-1.5 hours, with an average of 50ul of fluid collected. Participants who were menstruating were exempted from mucosal sampling. Urine (5ml) was collected for STI detection. Samples were stored at -80ºC for downstream analysis.

**HIV testing**

Rapid HIV testing was performed using a finger-prick blood sample, following Kenya’s National guidelines. Samples were tested with Shenghai Kehua Bioengineering (KHB) HIV (1+2) Rapid Diagnostic Kit. Participants who tested positive on the KHB kit then had confirmatory test using the First Response HIV 1-2.0 Rapid Whole Blood Test. HIV results were confirmed by qPCR of dry blood spots at the JC Wilt Infectious Disease Research Centre, Public Health Agency of Canada, Winnipeg, Canada.
Cytokine analysis
We quantified levels of 39 cytokines SoftCup™ supernatants diluted 1:10 using Biorad kits, according to manufacturer’s instructions, including (1) human Th17 panel (IL17F, GMCSF, IFNγ, IL10, MIP3α [CCL20], IL12P70, IL13, IL15, IL17A, IL22, IL9, IL1β, IL33, IL2, IL21, IL4, IL23, IL5, IL6, IL17E, IL27, IL31, TNFα, TNFβ, IL28α); (2) the human cytokine/chemokine panel I (IFNα2, IL12P40, IL1Ra, IL1α, IL7, IL8 [CXCL8], IP10 [CXCL10], MIP1α [CCL3], MIP1β [CCL4], RANTES [CCL5]); and (3) the human cytokine/chemokine panel III (MIG [CXCL9], ITAC [CXCL11], MIP3α [CCL20], MIP3β [CCL19], IFNλ1). Most (26/39) cytokines had >60% of specimens in the detectable range and were analysed as log10-transformed continuous variables. The remaining 13/39 cytokines (IL28α, TNFβ, IL21, IL27, IL17E, IL5, IL4, IL2, IL9, IL22, GMCSF, IL17F and CCL19), were analysed as binary variables. IL1Ra values were above the detection limit in >90% of samples and excluded from further analysis. The lower limit of quantification (LLOQ) was defined as the concentration in which the standard curve co-efficient of variation (CV) was >15%. All values below the LLOQ were imputed as the LOOQ/2. Intra-plate variability was determined using six sets of duplicate wells (CV range 9-11%). Inter-assay variation was assessed by running a common sample on each plate (CV range 7-17%).

Microbiome analysis
PCR amplification of the 16S rRNA gene 515f-806rB (V4) region was used to characterize the vaginal microbiome. DNA was extracted from SoftCup™ cell pellets using Qiagen DNeasy blood and tissue kits. Samples were amplified in triplicate and verified on agarose gels. 10ng of DNA/sample was pooled and amplicons cleaned using UltraClean PCR Clean-Up Kit. The Earth Microbiome sequencing Protocol (EMP) was followed. DNA was heat denatured into single strands to create a 2nM sample library, which was combined with a PhiX Control. The mixture was then loaded into the MiSeq v2 300 cycle reagent cartridge at 10pM for sequencing. Sample sequencing quality was ensured using FASTQC (V 0.11.5) and MultiQC (V 1.0) modules within the QIIME
toolbox (V 1.9.1)\textsuperscript{22}. Poor quality reads (Phred score<28) were then removed. Forward and reverse reads were joined using the PandaSeq (V 2.8) assembler program\textsuperscript{23}. The merged individual files were combined and closed reference OTUs picked using the Greengenes database (13_8 release)\textsuperscript{24}, with relative abundance reported at the genus and/or species level. Alpha diversity was calculated using the phylogenetic diversity whole tree method, and beta-diversity using UniFrac distances generated with jackknifed support using QIIME for phylogenetic-based analyses. Alpha and beta rarefaction was computed to a maximum sample depth of 2000. Additional sensitivity analyses were performed classifying OTUs using QIIME2 and the Silva database.

Detection of sexually transmitted infections (STIs)
Cryopreserved urine was tested for six STI-associated pathogens using the Anyplex II multiplex real-time PCR platform from Seegene (\textit{Chlamydia trachomatis}, \textit{Neisseria gonorrhoeae}, \textit{Trichomonas vaginalis}, \textit{Mycoplasma hominis} and \textit{genitalium}, and \textit{Ureaplasma urealyticum})\textsuperscript{25-27}.

Statistics
Baseline characteristics were compared using ANOVA and Chi-squared tests for continuous and categorical variables, respectively. Microbiome-related variables were defined in three main ways: 1) as categorical on the basis of unsupervised hierarchical clustering\textsuperscript{28}; 2) alpha diversity as a continuous variable, defined as above; and 3) beta diversity as continuous variables, including the first three principle coordinates in PCoA. Hierarchical clustering of microbiome profiles was carried out based on the Euclidean distance matrices, and the relative abundance of each genus within individual samples was generated using Ward’s linkage method, which was visualized in a heatmap generated using MetaboAnalyst 3.0\textsuperscript{29}. Spearman’s rank correlation was used to analyse associations between continuous microbiome, cytokine and sexual activity variables. In certain cases where the data were non-normally distributed, such as number of sex acts in the last week and time since last sex, data were categorized and analysed as predictors of alpha diversity in ANOVA.
including Dunn's post-hoc test. Multiple comparisons were adjusted using False Discovery Rate (FDR) method of Benjamini and Hochberg, with FDR (Q)=5%. We used multivariable linear regression analyses to determine associations of the vaginal microbiome and study group, modeled as three categories, with alpha diversity as the outcome variable. Covariates including age, condom and contraception use, HIV and STI status, number of sexual partners in the last week, and practicing vaginal douching were included in some models. These were pre-determined based on their potential to impact the mucosal milieu, and/or their association with study group. Statistical analyses were carried out using SPSS version 24 and 25. P values of p<0.05 (two-sided) were considered statistically significant.

Results

Participant characteristics

Study participants included self-identified FSW (n = 72), transactional sex (n = 30), and non-FSWs (n = 66, Table 1). The median ages of FSW, transactional, and non-sex worker participants were 20 (IQR: 18-22), 19 (IQR: 17-21), and 20 (IQR: 18-22) years, respectively (p=0.081). A subset of participants (18%) were <18 years of age. Approximately one-third of participants had completed primary school. The number of sexual partners differed by study group, with FSW, transactional, and non-sex workers having medians of 6 (IQR 3-11), 1 (IQR 0-4), and 0 (IQR 0-1) sexual partners last week, respectively (p<0.001). The prevalence of HIV among FSW was 12.5% (9/72), 6.6% among the transactional group (2/30) and 4.5% among the non-sex worker group (3/66). Similar proportions had ≥1 bacterial STI, with 43.0% (31/72), 30.0% (9/30), and 31.8% (21/66) for FSW, transactional, and non-sex workers, respectively. The most prevalent STIs were Mycoplasma hominis and Ureaplasma urealyticum at 21 and 22%, respectively.
Microbiome profiles including those associated with molecular-BV

16S rRNA sequencing was used to determine the relative abundance of cervicovaginal bacteria at the genus and/or species level. Unsupervised hierarchical clustering categorized the operational taxonomic units (OTUs) into five clusters (Figure 1A). Because these were highly similar to what was defined by Anahtar et al., we employed a similar naming system. *Lactobacillus* spp. were further differentiated into *L. iners* and *L. non-iners*, reflecting the biological differences within this genus. Cluster 1a (7%) was a mixture of *L. iners* and *L. non-iners* spp., while Cluster 1b (8%) was composed primarily of *L. non-iners*. Cluster 2, which accounted for 22% of the sample, was composed primarily of *L. iners*. In combination, Cluster 1a, 1b and 2 accounted for 37% of study participants (62/166). Bacterial communities associated with BV, including those dominated by *Gardenella* spp. (Cluster 3) and *Prevotella* spp. (Cluster 4) were characteristic of 63% (104/166) of participants. Principal coordinates analysis (PCoA) was used to compare beta diversity between clusters (Figure 1B). Consistent with unsupervised clustering, these data highlight the increased genetic diversity of Clusters 3 and 4. Similar clusters were observed in the analysis that made use of the Silva database (Supplemental Table 1 http://links.lww.com/QAI/B481).

Associations between microbiome clusters and cervicovaginal cytokine expression

We next carried out a principal component analysis on cytokine concentrations. A large proportion (60% by PC1 and PC2) of the variance in cytokine levels was explained by microbiome clustering (Figure 2A). Comparing cytokine expression between different microbiome clusters identified several significant associations, including IFN-α, CCL5, IL-7, IL-6, IL-1β, IL-1α, TNF-α, CXCL10, and CXCL11 (all p<0.001, ANOVA, Figure 2B). Most cytokines were present at higher concentrations in cluster 3 and 4, while chemokines CXCL10 and CXCL11 had the opposite trend. In support of this, we found modest correlations between alpha and beta diversity and higher cytokine concentrations: IL-10, IL-1α, IFN-γ, IL-13, IL-12p40, IL-7, IFN-α2, IL-15, IL-12p70, IL-1β, IL21, IL-23 and TNF-α (p<0.05, Supplemental Table 2 http://links.lww.com/QAI/B481). In
contrast, inverse correlations were observed between alpha and beta diversity and a number of chemokines, including CXCL10, CXCL11, CXCL9, CCL20, and CCL19 (p<0.05).

**Effect of age on microbiome diversity**

Age did not correlate strongly with microbiome cluster, although a trend was observed for cluster 1b being more common in younger participants (ANOVA p=0.053, Supplemental Figure 1A [http://links.lww.com/QAI/B481](http://links.lww.com/QAI/B481)). No association was observed between age and alpha diversity (beta = -0.012, 95% CI: -0.082 to 0.059, p=0.748, Table 2 and Supplemental Figure 1 [http://links.lww.com/QAI/B481](http://links.lww.com/QAI/B481)). Similarly, no associations were observed between age at sexual debut or years of sexual activity and any microbiome cluster. There was an inverse trend between alpha diversity and age of sexual debut (ANOVA, p=0.0726), and women who had been sexually active for 6-10 years had higher alpha diversity compared to women with 1-5 years of sexual activity (Supplemental Figure 1B [http://links.lww.com/QAI/B481](http://links.lww.com/QAI/B481)). These data suggest that age, age at debut, and duration of sexual activity make relatively minor contributions to vaginal microbiome composition in the young women enrolled in our study.

**Effect of demographic and behavioral characteristics on microbiome profile**

We next assessed whether self-identifying as an FSW was associated with vaginal microbiome alpha diversity in linear regression analysis (Table 2). Compared to the non-sex worker group, both the FSW and transactional group had higher alpha diversity. This was statistically significant for FSW vs non-FSW (beta=0.53, 95% CI: 0.16 to 0.90, p=0.01) but not non-FSW vs transactional (beta=0.28 95% CI: -0.20 to 0.76, p=0.257). The presence of an STI (p<0.001) and HIV infection (p=0.015) were also associated with alpha diversity. In multivariable linear regression analyses adjusted for age, DMPA use, unprotected sex, number of partners in the last week, vaginal douching, and HIV and STIs, alpha diversity remained significantly higher in the FSW group.
(beta=0.47, 95% CI: 0.05 to 0.90, p=0.03, Table 2). Similar results were obtained when the association between study group and alpha diversity was adjusted for time since last sex.

We carried out bivariable analyses to assess whether the association between FSW and alpha diversity may be mediated by frequency of sex and/or rate of partner change. The number of sex acts with all partners and the number of partners in the past week both correlated with alpha diversity across all groups (Spearman rank correlation $r=0.215$, $p=0.006$ for number of sex acts; $r=0.206$, $p=0.008$ for number of partners). We then compared alpha diversity with the number of sex acts within each study group (Figure 3A). In general, alpha diversity increased with the number of sex acts, but this was most evident in the non-FSW who had 0 compared to 1-10 sex acts in the last week (ANOVA $p=0.04$). We further assessed the role of condom use on microbiome diversity, stratified by group (Figure 3B). Similar alpha diversity was observed for “always”, “sometimes” and “never” condom users, although these participants generally all had higher alpha diversity than those who did not have any reported sex acts in the last week. In support of the finding that any sex in the past week was associated with alpha diversity, we observed a similar association between time since last sex and increased alpha diversity (ANOVA $p<0.001$, Figure 3C). These data suggest that both the amount and recentness of sexual activity may be upstream drivers of the association between the vaginal microbiome and engagement in sex work and/or transactional sex, leading to changes in the composition of vaginal microbial communities and resulting inflammatory profiles that have been associated with increased HIV acquisition risk.

**Discussion**

The cervicovaginal environment is an important determinant of HIV acquisition risk$^{30}$. In sub-Saharan Africa, where two-thirds of global HIV transmission occurs, young women and FSW represent two important key populations for focusing HIV prevention$^{17}$. Biological risk in these populations is incompletely understood, particularly in FSW who are new to sex work. Here we
present a detailed characterization of the HIV-risk associated vaginal micro-environment in young women, including those <18 years of age. Our results show that self-reported sex work and recent sexual activity correlate with increased vaginal microbiome diversity, which in turn explains a high proportion of the variance associated with cervicovaginal cytokine levels, including those associated with HIV acquisition. This is not to say other variables, including STIs, don’t also have strong cytokine effects. Similar to other studies, we observed reversed trends for pro-inflammatory cytokines such as IL-1α and IL-1β, and chemokines such as CXCL10, which were increased and reduced in women with molecular-BV, respectively. Many factors may influence cytokine levels in the female genital tract; however, our study, in agreement with others, suggests that the commensal bacteria in this mucosa represent a dominant influence on cytokine milieu.

Many studies have demonstrated that molecular- and/or Nugent-BV is common in young women of African descent. A small proportion (<20%) of our participants had Lactobacillus non-iners spp. thought to be “optimal” for protection against HIV. This is in line with a study of South African women where only 14% carried this “healthy” microbiome. Replication of this finding is important, demonstrating its generalizability to multiple regions of Africa. Whether this high rate of molecular-BV is due to differences in genetic, environmental and/or behavioral factors, including diet, antibiotic use and contraception methods, requires further research. The observation that BV has been associated with increased HIV risk and adverse reproductive outcomes underscore the need for HIV risk reduction efforts to include better methods to reduce inflammation and restore Lactobacillus dominant microbial communities in young women at high risk of HIV.

When certain microbial communities become established in the vaginal mucosa is difficult to ascertain, and this was not possible in the current cross-sectional study. However, the similar distribution of microbiome clusters and alpha diversity across all ages in our study (14–24 years) suggests that diverse vaginal microbiomes are common in young women under 18 years, in line
with suggestions that the vaginal microbiome is likely “set” very early in life during puberty\textsuperscript{36,37}. The extent to which a certain microbiome cluster is inherent to individual women remains unclear, but our data suggest that further investigation of host genetic associations with vaginal microbiota profiles is warranted\textsuperscript{38}.

Sexual activity is an important influence on the vaginal microbiome. Previous studies have shown that reporting a new sexual partner is associated with BV incidence\textsuperscript{39}. A study by Ravel et al that sampled women frequently showed that sexual activity was one of the few consistent associations of microbial diversity\textsuperscript{14}. In our study sex work was associated with molecular-BV; this could be due to increased sexual activity, as supported by recent sex correlating with alpha diversity. While the sex work-microbiome association was independent of recent sex, sexual exposure is likely an important driver of this association. This is supported by the fact that the recent sex-alpha diversity association was strongest in the non-sex worker group. This might indicate that very few sex encounters can increase the incidence of molecular-BV, with additional sexual encounters having minimal effects. Although we controlled for many potential confounders in multivariable analysis, it remains possible that factors in addition to sexual activity, such as the type of partners, condom use, and/or intravaginal practices, also contribute toward FSWs having increased alpha diversity. Our results are in agreement with one previous study that observed an association between the vaginal microbiome and sex work\textsuperscript{40}.

Our study had some limitations. Although our sample size was relatively small, we selected specimens randomly from a larger study that used population-representative sampling\textsuperscript{18}, increasing the generalizability of the findings. A larger study would be required to better understand how diversity of partner types (casual vs regular, paid vs unpaid, etc.), volume, and the importance of condom and/or semen exposure influence these microbiome associations. Sexual activity was self-reported via face-to-face interviews and therefore subject to social-desirability bias and under-
reporting. The intention of this analysis was on profiling the most abundant bacterial communities in our study population rather than a more detailed examination of bacterial function. However, for the vaginal microbiome, clear groupings by abundance category are evident, and are highly associated with immunological profiles including those that predict increased HIV acquisition risk\(^4,6\).

In conclusion, our data show that molecular-BV is common in young Kenyan women and may be an important driver of HIV risk-associated inflammation. Epidemiologically, this key population has amongst the highest HIV incidence rates in the region, emphasizing the need to better understand their risk profiles. Our study covers a wide range of sexual frequency, with many measures of sex work and/or recent sex correlated with molecular-BV. A key finding is that sex work exposure itself may increase the proportion of women with an HIV susceptible phenotype. This could be highly relevant to HIV transmission; if sex work induces microbiota changes, this could increase the likelihood of infection if the next sexual partner is HIV infected and viremic. These data further underscore the urgent need for better strategies to reduce the burden of BV, particularly in young women in eastern and southern Africa.

References


3. Masson L, Arnold KB, Little F, et al. Inflammatory cytokine biomarkers to identify women with asymptomatic sexually transmitted infections and bacterial vaginosis who are at high


**Figure legends**

**Figure 1.** Cervicovaginal microbiota profiles in adolescent girls and young women from Mombasa, Kenya as determined by 16S rRNA sequencing of SoftCup pellets (n = 162). **A**) Unsupervised clustering of operational taxonomic unit (OUT) abundance into microbiome cluster. The heatmap is colour-coded by % abundance, with the name of each cluster summarized on the top of the graph. **B**) Principle coordinates analysis (PCoA) using unifrac distances between individuals. The colours in B) correspond to the clusters indicated in A), with communities dominated by *Prevotella* spp. (cyan bar/dots), *Gardenella* spp. (pink bar/dots), *Lactobacillus iners* (blue bar/dots), *Lactobacillus non-iners* (green bar/dots), and mixed *Lactobacillus* spp. (red bar/dots).

**Figure 2.** Relationship between vaginal microbiota and cervicovaginal cytokine/chemokine concentrations. **A**) Principal component (PC) analysis showing unsupervised clustering of cervicovaginal (Softcup) cytokine/chemokine variation by microbiome cluster. Each dot represents an individual colour-coded by cluster from Figure 1 and Figure 2A. The % variance captured is shown on the axes for each principle component. **B**) Chemokine and cytokine concentrations in Softcup fluid from individuals grouped by microbiome cluster. Lines indicate the median and interquartile range (IQR) for each dataset. Groups were compared using Kruskal-Wallis test with Dunn’s post-hoc analyses.

**Figure 3.** Association between recent sexual activity and alpha diversity. **A**) Differences in alpha diversity within study groups (non-sex worker, transactional and FSW) based on number of sex acts in the preceding week. Number of sex acts are either 0 if no sex in the last week or grouped 1-10, 11-20, 21-30 and ≥30. **B**) Differences in alpha diversity within study groups based on condom use pattern in the preceding week, coded as 0 if no sex, and in categories of always, never, and sometimes using condoms. **C**) Differences in alpha diversity based on time since last sex, categorized as < 1, 1-7, 8-15, 16-30, and >30 days since last reported sex act. Lines indicate the median and interquartile range (IQR) for each dataset. Groups were compared using Kruskal-Wallis test with Dunn’s post-hoc analyses.
<table>
<thead>
<tr>
<th></th>
<th>Non-FSW</th>
<th>Transactional</th>
<th>FSW</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>66</td>
<td>30</td>
<td>72</td>
<td>-</td>
</tr>
<tr>
<td>Age, median (IQR)</td>
<td>20 (18,22)</td>
<td>19 (17,21)</td>
<td>20 (18,22)</td>
<td>-</td>
</tr>
<tr>
<td>HIV status (n/N)</td>
<td>4.5% (3/66)</td>
<td>6.6% (2/30)</td>
<td>12.5% (9/72)</td>
<td>-</td>
</tr>
<tr>
<td>Completed primary school (n/N)</td>
<td>31.8% (21/66)</td>
<td>40.0% (12/30)</td>
<td>31.9% (23/72)</td>
<td>-</td>
</tr>
<tr>
<td>Number of partners in the last week of sex (n/N)</td>
<td>0 (0,1)</td>
<td>1 (0,4)</td>
<td>6 (3,11)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Vaginal sex acts in the last week, median (IQR)</td>
<td>0 (0,1)</td>
<td>2 (0,6)</td>
<td>7 (3,17)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Time since last sex in days median (IQR)</td>
<td>15 (4, 60)</td>
<td>9 (1,30)</td>
<td>2 (1,4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>STI positive (n/N)</td>
<td>31.8% (21/66)</td>
<td>30.0% (9/30)</td>
<td>43.1% (31/72)</td>
<td>-</td>
</tr>
<tr>
<td>DMPA use</td>
<td>12.1% (8/66)</td>
<td>20.0% (6/30)</td>
<td>23.6% (17/72)</td>
<td>-</td>
</tr>
</tbody>
</table>

FSW: female sex worker; IQR: interquartile range; DMPA: depo-medroxyprogesterone acetate; STI: sexually transmitted infection; HIV: human immunodeficiency virus
Table 2. Multivariable linear regression analysis of variables associated with vaginal microbiota alpha diversity

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Univariate analysis</th>
<th></th>
<th></th>
<th>Multivariable analysis</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (95% CI)</td>
<td>P value</td>
<td>Estimate (95% CI)</td>
<td>P value</td>
<td></td>
</tr>
<tr>
<td>Female sex worker</td>
<td>0.53 (0.16 – 0.90)</td>
<td>0.010</td>
<td>0.47 (0.05 – 0.90)</td>
<td>0.030</td>
<td></td>
</tr>
<tr>
<td>Transactional</td>
<td>0.28 (-0.20 – 0.76)</td>
<td>0.257</td>
<td>0.26 (-0.22 – 0.74)</td>
<td>0.288</td>
<td></td>
</tr>
<tr>
<td>Non-sex worker</td>
<td>Ref</td>
<td></td>
<td>Ref</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.01 (-0.08 - 0.06)</td>
<td>0.748</td>
<td>-0.04 (-0.11 – 0.03)</td>
<td>0.254</td>
<td></td>
</tr>
<tr>
<td>Any unprotected sex</td>
<td>-0.02 (-0.11 – 0.07)</td>
<td>0.661</td>
<td>-0.05 (-0.15 - 0.06)</td>
<td>0.352</td>
<td></td>
</tr>
<tr>
<td>DMPA use</td>
<td>0.09 (-0.34 – 0.52)</td>
<td>0.677</td>
<td>0.01 (-0.44 – 0.46)</td>
<td>0.962</td>
<td></td>
</tr>
<tr>
<td>HIV positive</td>
<td>0.79 (0.14 – 1.44)</td>
<td>0.018</td>
<td>0.77 (0.15 – 1.39)</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>STI positive</td>
<td>0.79 (0.45 – 1.12)</td>
<td>&lt;0.001</td>
<td>0.80 (0.46 – 1.14)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Number of sex acts last week</td>
<td>0.00 (-0.02 – 0.02)</td>
<td>0.942</td>
<td>-0.01 (-0.04 – 0.03)</td>
<td>0.687</td>
<td></td>
</tr>
<tr>
<td>Practices vaginal douching</td>
<td>0.18 (-0.16 – 0.52)</td>
<td>0.306</td>
<td>0.15 (-0.19 – 0.49)</td>
<td>0.379</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1
Figure 2
Figure 2 (con't)
Figure 3

α-diversity

Non sex worker
Transactional
Female sex worker

Number of sex acts in the last week

Condom use in the last week

0= No sex
1= Always use condoms
2= Never use condoms
3= Use condoms sometimes

Figure 3
Figure 3 (con’t)