Associations Between Vaginal Microbiota and Sexual Activity in Kenyan Adolescent Girls: A Focus on *Lactobacillus* and *Limosilactobacillus* Species

Molly Sneller

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington
2025

Committee:
Alison Roxby
Lisa Manhart
Nelly Mugo

Program Authorized to Offer Degree:
Epidemiology

University of Washington

Abstract

Associations Between Vaginal Microbiota and Sexual Activity in Kenyan Adolescent Girls: A Focus on *Lactobacillus* and *Limosilactobacillus* Species

Molly Sneller

Chair of Supervisory Committee:

Alison Roxby

Department of Epidemiology

Background: The vaginal microbiota is a key component of genital and reproductive health. Sexual activity can influence the dominant microbes present, shifting away from colonization with key protective bacteria. This study characterized the relationship between vaginal *Lactobacillus* and *Limosilactobacillus* species and history of sexual activity.

Methods: This cross-sectional study used data collected at the baseline visit for the Kenya Girls Study cohort of adolescent girls and young women (AGYW) aged 16–18 in Thika, Kenya. Bacterial composition was determined from vaginal swabs using 16S rRNA gene sequencing. Participants were classified as sexually active based on self-report and biological markers (incident STI, Y chromosome detection, spermatozoa on Gram stain, or pregnancy). Microbial diversity was evaluated using Shannon diversity indices and Bray-Curtis distances. Poisson regression with robust standard errors was used to estimate associations between sexual activity and detection of specific *Lactobacillus* and *Limosilactobacillus* species. A compositional LASSO model incorporating relative abundance data was used to describe odds of sexual activity, and an additional analysis assessed correlates of *Limosilactobacillus* species detection.

Results: We included 142 AGYW, 65.5% who had never had sex and 34.5% who were classified as sexually active. The most common species detected by relative abundance was *Lactobacillus crispatus/gallinarum.* Beta diversity analysis using Bray-Curtis distances revealed significant differences in community composition between groups (PERMANOVA $p = 0.009$, $R^2 = 0.0332$, while *Lactobacillus iners* was more often detected in the sexually active group [PR unweighted 1.36 (1.02-1.83) $p= 0.039$, PR weighted 1.49 (0.87-2.55) $p = 0.148$], and detection of *Lactobacillus jensenii/mulieris* was also associated with sexual activity [PR unweighted 2.44 (1.12-5.29) $p=0.024$, PR weighted 3.09 (1.27–7.52) $p = 0.014$]. A compositional LASSO model showed that odds of belonging to the sexually active group increased with higher relative abundance of *L. iners* (OR 1.022) and decreased with higher abundance of *L. crispatus/gallinarum* (OR 0.978). *Limosilactobacillus* detection was positively associated with detections of *L. crispatus/gallinarum, L. iners,* and *L. gasseri/paragasseri.*

Conclusion: This study found that sexual activity among AGYW in Kenya was associated with increased detection of *L. iners*. Understanding functional roles of diverse Lactobacilli could inform interventions to promote an optimal vaginal microbiome.

Introduction

Adolescent girls and young women (AGYW) account for a high proportion of the global population at risk for acquiring sexually transmitted infections (STIs) and human immunodeficiency virus (HIV), with high rates of infection in sub-Saharan Africa.[1] A proposed mechanism for the increased susceptibility to infection is that mucosal barriers are affected by the stability and diversity of the vaginal bacterial microenvironment, mediated by shifts in the vaginal pH and increased local inflammation.[2,3] Bacterial vaginosis (BV), a dysbiotic state of the vaginal microbiota characterized by dominance of many non-*Lactobacillus* bacteria, including anaerobic gram-negative bacteria such as *Gardnerella* and *Prevotella* species, is associated with a variety of negative reproductive health outcomes such as cervicitis and increased rates of STIs across multiple populations of cisgender women and girls.[4] BV is also more common among individuals that have had prior or recent sexual activity.[5–9] Conversely, the dominance of certain Lactobacilli*,* such as *Lactobacillus crispatus* and *Lactobacillus jensenii,* in the vaginal microbiota has been previously associated with a lower risk of bacterial STIs including gonorrhea and chlamydia; viral STIs including HIV, HPV and HSV-2; and adverse pregnancy outcomes.[10–15]

While *Lactobacillus*-dominated vaginal microbiomes are often described as “optimal”, many AGYW with optimal microbiomes exhibit variation in the specific *Lactobacillus* species present and the type of Lactobacilli which are present may confer differing levels of protection. *Lactobacillus iners*, for example, one of the most common species found in the vaginal microbiota of young women in sub-Saharan Africa, has been previously described as an indicator of a sub-optimal state of the vaginal microbiome. It is also more frequently present in sexually active individuals and associated with increased risk of STI.[16–19] Grouping all Lactobacilli together, a common research practice if using Gram stain results, can therefore obscure the effects of different *Lactobacillus* species. Low population prevalences of *L. crispatus* in some areas have also motivated in some studies to group *L. crispatus* and *L. iners*. Given the differences in outcomes of women colonized by different *Lactobacillus* species, collapsing the vaginal microbiota into *Lactobacillus*-dominant and therefore "optimal" or non-*Lactobacillus*-dominant or "sub-optimal" is not ideal.

Limosilactobacillus spp. are lactic acid-producing bacteria that, based on preliminary data, are in both high abundance and prevalence in this cohort. However, our understanding of the role of *Limosilactobacillus spp.* in the vaginal microenvironment is incomplete. Multiple current species of *Limosilactobacillus* were previously classified as Lactobacilli*,* including *Limosilactobacillus reuteri* (basonym: *Lactobacillus reuteri*) and *Limosilactobacillus vaginalis* (basonym: *Lactobacillus vaginalis).*[20] Prior research has noted high co-occurrence of *Limosilactobacillus* species with *L. crispatus* and *L. jensenii* in the vaginal microbiota,[6] suggesting that these species exist with other beneficial lactic-acid producing bacteria.

To provide more specific resolution about the different Lactobacilli in young women, we studied a cohort of AGYW in Kenya who participated in a reproductive health study. AGYW in the cohort were found to have a low prevalence of BV and high prevalence of Lactobacilli. Our goal was to determine whether a history of sexual activity was associated with differences in the relative abundance of different Lactobacilli among AGYW to understand how the vaginal microbiota may by influenced by early sexual life and to further define protective factors in vaginal microbiota. We also investigated correlates of *Limosilactobacillus* detection.

Methods

Study design and setting

This study is a cross-sectional retrospective analysis conducted at point of entry into the Kenya Girls Study (KGS) longitudinal cohort. In the original cohort, adolescents and young women were followed over 5 years to evaluate STIs and vaginal inflammation during the years spanning first sex. This study used data collected from participants at their first clinical research site visit in Thika, Kenya, a peri-urban area located 35 kilometers northeast of Nairobi with a population of approximately 250,000 people. The clinical research clinic provides onsite care (STI, family planning, menstrual health) to study participants.

Study participants

Recruitment for the original cohort occurred between 2014-2016 through community-based methods.[21] Study participants include 400 HIV-seronegative and HSV-2 seronegative adolescent girls ages 16-18 years. The original criteria for inclusion in the cohort was no history of sexual intercourse or sexual experience with only one lifetime partner. At enrollment, about 20% of participants were classified as having had vaginal sex by self-report or by positive biological marker of sex including incident STI, Y chromosome detection, spermatozoa on Gram stain, or pregnancy. For the parent study, first-time STI cases were matched using incidence density sampling at a 1:1 ratio with a sexually active STI-negative control and a pre-sex STI-negative control. All participants completed a sexual history questionnaire at their first clinic visit. Microbiome sequencing on the first visit vaginal swab was only performed on participants identified as cases or controls. This study used all cohort participants who had a sequenced vaginal microbiome specimen from the study enrollment visit.

Data collection & Laboratory analysis

Nugent score was determined via Gram staining of vaginal smears. Dacron vaginal swabs were collected and cryopreserved at -80°C for vaginal microbiome processing at the University of

Calgary's Western Canadian Microbiome Centre. DNA was extracted from vaginal swabs, followed by broad-range 16S rRNA gene PCR and high-throughput sequencing (Ilumina MiSeq) with a two-amplicon approach, primers 319F (ACACTGACGACATGGTTCTACA) and 806R (TACGGTAGCAGAGACTTGGTCT), to accurately classify bacterial species and assess relative abundance. Sequencing data were processed using the DADA2 pipeline in R, including quality filtering, chimera removal, and taxonomic classification using reference databases.[22–24] Amplicon sequence variant (ASV) counts were collapsed for some species groups including (*L. crispatus* and *L. crispatus/gallinarum), (L. gasseri, L. paragasseri,* and *L. gasseri/paragasseri*), and (*L. jensenii* and *L. mulieris*) across all analysis.

Statistical analysis

The exposure in this analysis was sexual activity by self-report or positive biomarker result. Although the analysis was limited to enrollment data, specimens were selected for broad range PCR and sequencing using incidence density sampling. All participants identified as incident STI cases during the longitudinal cohort were included in microbiome sequencing analysis. At each study visit where a case was identified, two controls were randomly selected from among all participants who were STI-negative and still at risk at that same visit; these control participants also had vaginal microbiome sequencing. We calculated inverse probability weights (IPWs) to account for the original sampling design and ensure representativeness of the cohort in cross-sectional analyses. The selection probability for each case was 1.0 for an IPW of 1.0; the selection probability for each control was calculated as the ratio of selected controls to all eligible controls at that visit, and the corresponding IPWs were the inverse of these probabilities.

Participant characteristics were compared between participants who were never sexually active and those with previous sexual activity, including age at study entry (years), education (years of schooling), monthly income (Kenya Shillings), residence (urban or rural), housing condition (High housing condition defined by formal settlement, running water, concrete floors, electricity, and roof made of tile or metal, Medium housing condition with at least 3 of these characteristics, and Low housing condition defined as 2 or fewer reported characteristics), age at menarche (years), and Nugent score, a clinical marker of BV (score of 7-10). Stacked bar charts were used to visualize relative abundance of vaginal bacterial species in each group. Alpha diversity for the never sexually active and previously sexually active groups was described using Shannon diversity indices and statistical significance determined by a Wilcoxon Rank-Sum test. Beta diversity was computed using Bray-Curtis distances to measure difference in microbial composition between these two groups; statistical significance of the differences was calculated using PERMANOVA and visualized using Principal Coordinate Analysis (PCoA).

We used Poisson regression with robust standard errors to estimate prevalence ratios (PR) with 95% confidence intervals (CI) to assess the association between history of sexual activity (exposure) and detection of select *Lactobacillus* and *Limosilactobacillus* species (outcome) in the vaginal microbiome samples. Regression models were adjusted for characteristics identified as potential confounders *a priori,* including age, housing condition and rural/urban residence.

A compositional LASSO approach was used to incorporate relative abundance of bacterial species in analyses of their association with sexual history status and to address collinearity among bacterial species. Odds ratios (ORs) computed by the LASSO model represent odds of belonging to the sexually active group per log2 increase in relative abundance of bacterial species. Zero ASVs counts in were replaced with a pseudocount of 0.5 prior to calculation of relative abundance and inclusion in the model.

In a secondary analysis we assessed correlates of detection of two *Limosilactobacillus* species (*L. coleohominis* and *L. reuteri/vaginalis*) using a multivariable Poisson regression with robust standard errors. Correlates evaluated included other high-abundance vaginal microbial species, sexual history, Nugent score, and demographic and socioeconomic factors such as age, education, and residence.

All analysis were done using R version 4. 4. 3.

Research Ethics

The study was reviewed by the Kenya Medical Research Institute Scientific Ethics Review Unit, the University of Washington Institutional Review Board, and the Calgary Conjoint Health Research Ethics Board. Written informed consent was obtained for participants over the age of 18. For participants under age 18, a parent/guardian provided written informed consent, and written participant assent was obtained separately from parents/guardians, to ensure assent was free from parental influence.

Results

Participant characteristics

At enrollment, 142 AGYW had vaginal swab specimens with microbiome testing results and were included in this analysis. Of these AGYW, 93 (65.5%) had no prior history of vaginal sex and 49 (34.5%) were classified as having previously been sexually active based on self-report in the administered questionnaire or a positive biomarker result. The median age of AGYW with sexual activity was higher (19.4 years [IQR 18.5 – 20.0]) than those that have never had sex (18.3 years [IQR 17.6 - 19.1]) (Table 1). Participants in both groups had similar median years of education (12 years) as well as little to no monthly income (ranged from 0 – 2000 Kenya

Shillings (KSh) = 0 – 10 USD). Also similar between groups was median age at menarche (14.0 years) (Table 1). It was more common for AGYW who were never sexually active to reside in rural setting (64.5%) compared to AGYW that were sexually active (49.0%).

AGYW were asked at the first visit if they had ever been sexually active; among the girls that were classified as never having had sex, 100% self-reported "No", while among the girls that were classified as sexually active, only 67.3% reported "Yes", meaning 16 (32.7%) individuals were identified as sexually active based on a positive biomarker (Table 1).

Vaginal Microbiome Diversity

We compared vaginal microbial diversity among AGYW by sexual activity status measured by the Shannon diversity index (Figure 1). In unweighted analyses, the mean Shannon index was 0.789 among those who reported never having had sex and 0.636 among those who had prior sex. However, the difference was not statistically significant (p = 0.201). These results did not differ after inverse selection probability weighting (0.617 vs. 0.612, p = 0.975).

Most participants in both groups had Nugent scores indicating a normal vaginal microbiota (0–3): (82) 88.2% among those who never had sex and (44) 89.8% among those who were sexually active. Nugent scores 7-10 were uncommon in both groups (Table 1). Relative abundance of taxonomic species differed across vaginal microbiome samples, however most vaginal swab specimens had highest relative abundance of *Lactobacillus crispatus/gallinarum* or *Lactobacillus iners* (Figure 2). Unsupervised principal coordinates analysis (PCoA) using Bray-Curtis distances revealed that Axis 1 accounted for 57.7% of variation and Axis 2 for 15.7% (Figure 3). Higher values on Axis 1 correspond to more *Lactobacillus iners* and less *Lactobacillus crispatus/gallinarum* whereas higher values on Axis 2 correspond to higher amounts of both species. To further assess differences in community composition, a PERMANOVA confirmed a significant difference in Bray-Curtis distances between participants who were sexually active and those who had not been sexually active ($p = 0.009$, $R^2 = 0.0332$).

Associations with detection of bacterial species

The presence of *Lactobacillus* species varied somewhat by sexual activity status (Table 2a). *Lactobacillus crispatus/gallinarum* was detected more frequently among participants who had never had sex compared to those who had in both the weighted (78.4% vs. 67.9%) and unweighted (71.0% vs. 55.1%) analysis, though not significant, whereas *Lactobacillus iners* was more often detected among previously sexually active AGYW (71.4% vs. 53.8% Unweighted; 60.5% vs 52.8% Weighted). Detection of *Lactobacillus gasseri/paragasseri* was similar between groups. For the most common *Limosilactobacillus* species, including *Limosilactobacillus reuteri/vaginalis* and *Limosilactobacillus coleohominis*, there was no significant difference in detection by group (Table 2a).

In Poisson regression models to assess associations between bacterial species and sexual history (Table 2a and 2b), after adjusting for age, housing condition, and rural/urban residence, detection of *L. iners* and *L. jensenii/mulieris* were significantly associated with sexual activity in the unweighted analysis (PR = 1.36 [95% confidence interval 1.02–1.83], $p = 0.04$ and PR = 2.44 [1.12-5.29], $p = 0.02$, respectively). Weighting attenuated the significance of the association for *L. iners* (PR = 1.49 [0.87-2.55], $p = 0.15$), but strengthened the association for *L. jensenii/mulieris.* Adjusted Poisson regression models showed an inverse association between *Lactobacillus crispatus/gallinarum* detection and sexual activity in both the unweighted and weighted models (PR = 0.77 [0.57-1.05], $p = 0.10$ and PR = 0.77 [0.54-1.10], $p = 0.16$), but this was not statistically significant (Table 2b). Compositional LASSO also revealed similar trends: the odds of belonging to the previously sexually active group were decreased per log2 increase in relative abundance of *Lactobacillus crispatus/gallinarum* (OR 0.978), and consistent with results from the detection models, odds of belonging to the sexually active group increased per log2 increase in relative abundance of *Lactobacillus iners* (OR 1.022) (Table 3).

Correlates of Limosilactobacillus Detection

In a multivariate Poisson Regression model evaluating correlates of *Limosilactobacillus* species detection, no significant associations were observed with demographic, behavioral or socioeconomic characteristics, including age, education, residence, or sexual activity adjusting for other correlates in the model (Table 4a). Higher Nugent score, however, showed an inverse association. Participants with BV-range scores (7–10) had lower likelihood of *Limosilactobacillus* detection in the weighted model (PR = 0.19 [0.05-0.81], $p = 0.03$). In the model assessing associations between *Limosilactobacillus* and other microbial species (Table 4b), *Lactobacillus crispatus/gallinarum* (PR = 1.50 [0.96-2.35], $p = 0.08$), *L. iners* (PR = 1.34 [0.99-1.81], $p = 0.06$), and *L. gasseri/paragasseri* (PR = 1.29 [0.99-1.67], $p = 0.06$) showed positive associations with *Limosilactobacillus* detection in weighted models, though not statistically significant when adjusting for other Lactobacilli.

Discussion

In this cross-sectional analysis comparing AGYW who had and had not initiated sexual activity in Thika, Kenya, we identified several key differences in population demographic characteristics and participants' vaginal microbiota. First, BV was rare in this population for both groups, meaning the vast majority of vaginal swab specimens had *Lactobacillus*-dominant vaginal microbiota based on Nugent score. Our results demonstrated that detection of *Lactobacillus crispatus* tended to be less likely in AGYW that were sexually active. Conversely, we were more likely to detect *Lactobacillus iners* and *Lactobacillus jensenii/mulieris* among sexually active AGYW.

We found that *L. crispatus* was less likely to be detected among AGYW who had prior initiation of sexual activity, and this was also true with increasing relative abundance of *L. crispatus.* This aligns with previous understanding that *L. crispatus* is a stable and protective member of the vaginal microbiota that is more often found in individuals with little to no recent sexual experience.[25,26] *L. iners* and *L. jensenii/mulieris,* on the other hand, were more often detected among sexually active participants which may reflect a shift in the vaginal microbiota that occurs with sexual activity. Some previous studies have shown a similar association for *L. iners* detection or high relative abundance in women who reported one or more sexual partners in the past year[27] as well as one study among Kenyan secondary school aged girls that observed increased odds of *L. iners* or *Gardnerella*–dominant vaginal microbiomes among previously sexually active girls compared to those that had never had sex.[28] It was surprising to observe higher likelihood of detecting *L. jensenii/mulieris* in the previously sexually active participants in our study, as this species has more frequently been grouped with *L. crispatus,* which is thought to confer protection against disturbances to the vaginal microbiota.[29] Overall detection was low for *L. jensenii/mulieris* and it was rarely the dominant Lactobacilli in a sample by relative abundance.

This research supports the idea that the ‘optimal’ vaginal microbiota is more complex than simply being dominated by any *Lactobacillus* species. Not all *Lactobacillus* species are equally protective against infections and species such as *L. iners* may be indicative of a sub-optimal microenvironment. One key hypothesis for the protective nature of a *Lactobacillus*-dominant microbiome is the ability of *Lactobacillus* species to produce lactic acid, which lowers the vaginal pH and inhibits colonization by other bacteria.[30] However, not all Lactobacilli equivalently produce lactic acid. In axenic environments, *L. iners* only produces one of two lactic acid isomers (L-lactic acid) while *L. crispatus* and *L. gasseri* are able to produce both L- and D-lactic acid.[31] The ability of *L. crispatus* to inhibit the establishment of other bacteria by creating a strongly acidic environment is important for the maintenance of a stable vaginal microbiome. Differences in lactic acid production underscore that *Lactobacillus spp.* are not all metabolically equivalent. Recognizing the distinct roles individual species may play in the vaginal microenvironment can help to understand how some Lactobacilli contribute to a stable and protective vaginal microbiome, while others may confer a more vulnerable microbial state.

In another study observing the influence of penile-vaginal sex on the vaginal microbiota, *L. iners* also showed higher detection in previously sexually active individuals, more similar to *Gardnerella vaginalis* than to *L. crispatus.*[19] Understanding how sexual activity influences the composition of Lactobacilli in vaginal microbiota and therefore its ability to confer protection from adverse sexual health outcomes can eventually be used to better inform health interventions for vaginal dysbiosis.

In the exploratory analysis for characterization of correlates of *Limosilactobacillus* species detection, we found that presence of these species was inversely correlated with BV. Further, we identified several bacterial species that co-occur with *Limosilactobacillus* including, *L. crispatus, L. iners,* and *L. gasseri/paragasseri* placing the genus as more often present in the normal vaginal microbiota.

Several previous studies have explored the presence of *Limosilactobacillus vaginalis* across various populations prior to its taxonomic nomenclature change using its basonym '*Lactobacillus vaginalis'*. [8,26,32] In one study comparing presence of bacterial species among populations of Kenyan, South African and Rwandan women, *L. vaginalis* was present in 48% of the Kenyan adolescent samples.[8] This was slightly lower than the proportion of specimens in which our study detected *L. reuteri/vaginalis*. In that study, *L. vaginalis* had a strong association with optimal Nugent scores of 0-3 in adult women, which was in agreement with our finding in which detection of *Limosilactobacillus* species was less likely in AGYW with BV (7-10) compared to those with optimal Nugent scores (0-3). While we found no significant difference in the presence of *L. vaginalis* between AGYW who previously had sex and those who did not, a study of girls aged 17-18 in Tanzania found that there was higher prevalence of *L. vaginalis* among girls that reported never having penile-vaginal sex (74%) compared to those who did report sex (55%).[26] The positive associations we found of *Limosilactobacillus* detection with detection of *L. crispatus, L. iners,* and *L. gasseri/paragasseri* place Limosilactobacilli as often detected with other lactic-acid producing members of the vaginal bacterial community.[33] One study that also found relatively high prevalence of *Limosilactobacillus* species in a cohort of women and girls in Belgium reported similar co-occurrence patterns with *Lactobacillus* species, specifically frequent co-detection of *L. crispatus* and *L. jensenii.*[6] Taken together, these findings suggest that the presence of *Limosilactobacillus* species may serve as a marker of a more optimal vaginal microbiota, given their reduced detection among individuals with elevated Nugent scores and co-occurrence with other *Lactobacillus* species.

A strength of this study is that species-level microbial identification was used to describe participants' vaginal microbiota, which enabled us to distinguish among specific *Lactobacillus* and *Limosilactobacillus* species. This level of taxonomic resolution is important as the concept of an "optimal" vaginal microbiota is more complex than being simply *Lactobacillus*-dominant. By capturing these differences on the species level, our study built a more robust picture of the complexities in the vaginal microbiota in adolescence and young adulthood through the key event of onset of sexual activity. Another strength is the study population which included adolescent girls and young women in Kenya prior to sexual activity. In previous vaginal microbiome research, highly studied populations are typically sexually active women; studies in adolescents often prioritize individuals that demonstrate 'high-risk' sexual behaviors. Lastly, methodologically this study incorporated inverse probability weighting to account for the case-control sampling design of the parent cohort. This reduced concern that cases of laboratory detected STI in the parent cohort were overrepresented in the microbiome analysis.

Limitations of this study include that it is observational and cross-sectional by design, and conclusions regarding the temporality of the association between sexual activity and microbiome composition cannot be made. Our results capture the differences between populations that were or were not sexually active at one point in time but cannot assess the durability of these changes over time. Some prior literature has observed variability in the vaginal microbiota across the menstrual cycle, data on menstrual cycle was not captured at the enrollment visit in our study.[34,35] Further we cannot infer the onset of sexual activity was the cause of any differences in the vaginal microbiota. There also remains potential for residual confounding by characteristics that could differ between the sexually active and never sexually active participants. Factors including antibiotic use, hormonal contraceptive use, and vaginal washing practices may influence the composition of the vaginal microbiota and were not measured at this study visit. Lastly, the main exposure of this study is history of sexual activity defined by self-report or biomarkers of vaginal intercourse. Due to the nature of self-reported data, there is opportunity for recall bias or social desirability bias. To help account for this, other information was used to define the exposure including incident STI, Y chromosome detection, spermatozoa on Gram stain, or pregnancy.

Future directions for this research could incorporate metagenomic or metatranscriptomic methods to further contextualize the functional activity of the bacteria in this environment to explain some of the differences seen in this study. We conclude that initiation of sexual activity is likely a critical perturbation event that may alter the vaginal microbiota composition in a way that shifts from the protective *L. crispatus* dominance to less optimal *L. iners.* Functional characterization of these bacteria to build a detailed metabolite profile of the varying protective properties of Lactobacilli and Limosilactobacilli could create new avenues for promoting optimal vaginal health.

Table 1. Demographic, socioeconomic, and health characteristics among Kenyan adolescent girls and young women by sexual activity at study entry.

Participant Characteristics	**Sexually Activity (Self-report or Positive Biomarker*)**	
	Never sexually active ***N = 93*** ***Median (IQR) or n (%)***	**Sexually active** ***N = 49*** ***Median (IQR) or n (%)***
Age (years)	18.3 (17.6 - 19.1)	19.4 (18.5 – 20.0)
Education (years)	12 (10 – 12)	12 (12 – 13)
Monthly income (KSh)**	0 (0 – 1000)	0 (0 – 2000)
Residence		
Rural	60 (64.5)	24 (49.0)
Urban	33 (35.1)	25 (51.0)
Housing condition †		
High	30 (32.3)	17 (34.7)
Medium	20 (21.5)	10 (20.4)
Low	43 (46.2)	22 (44.9)
Age at menarche (years) ††	14.0 (13.0 – 15.8)	14.0 (13.0 – 15.0)
Self-report of sexual activity		
Yes	0 (0)	33 (67.3)
No	93 (100)	16 (32.7)
Nugent Score §		
0-3	82 (88.2)	44 (89.8)
4-6	5 (5.4)	3 (6.1)
7-10	6 (6.5)	2 (4.1)

---

* Positive biomarker result was determined by presence of pregnancy, STI (gonorrhea, chlamydia, trichomonas, and herpes simplex virus 2), spermatozoa on Gram stain, or detection of Y chromosome

** 1000 KSh ~ 10 USD

† High housing condition was defined as reported formal settlement, running water, concrete floors, electricity, and roof made of tile or metal. Medium condition was defined as at least 3 of the 5 above characteristics. Low condition was defined as 2 or less of the reported characteristics.

†† Age of menarche had missing information for 10 individuals in the never sexually active group and 8 individuals in the sexually active group.

§ Nugent score of ≥7 was diagnostic of Bacterial Vaginosis (BV)

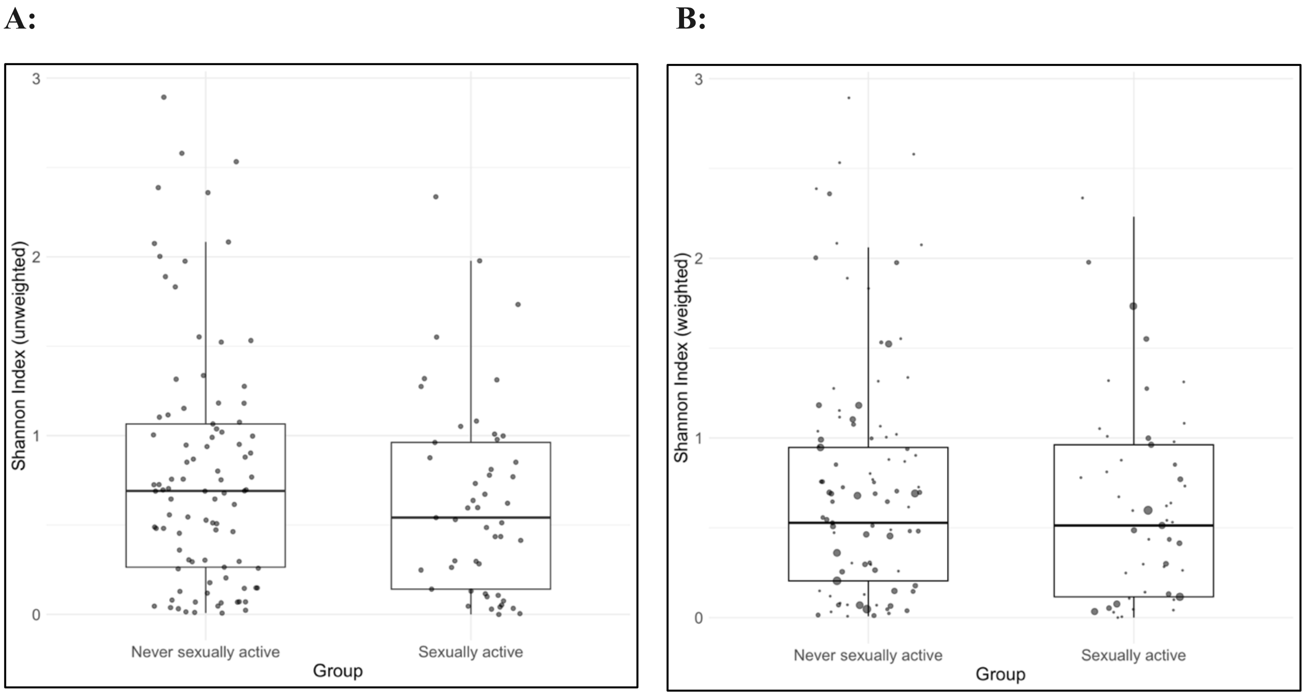

Figure 1. Vaginal microbiome diversity using Shannon Diversity Index stratified by sexual activity (A) Unweighted vs (B) Weighted by inverse probability of selection. The size of each overlaid point its proportional to its weight in the analysis. Boxes represent the interquartile range (IQR) with the median indicated by a horizontal line and whiskers extending to 1.5 x IQR.

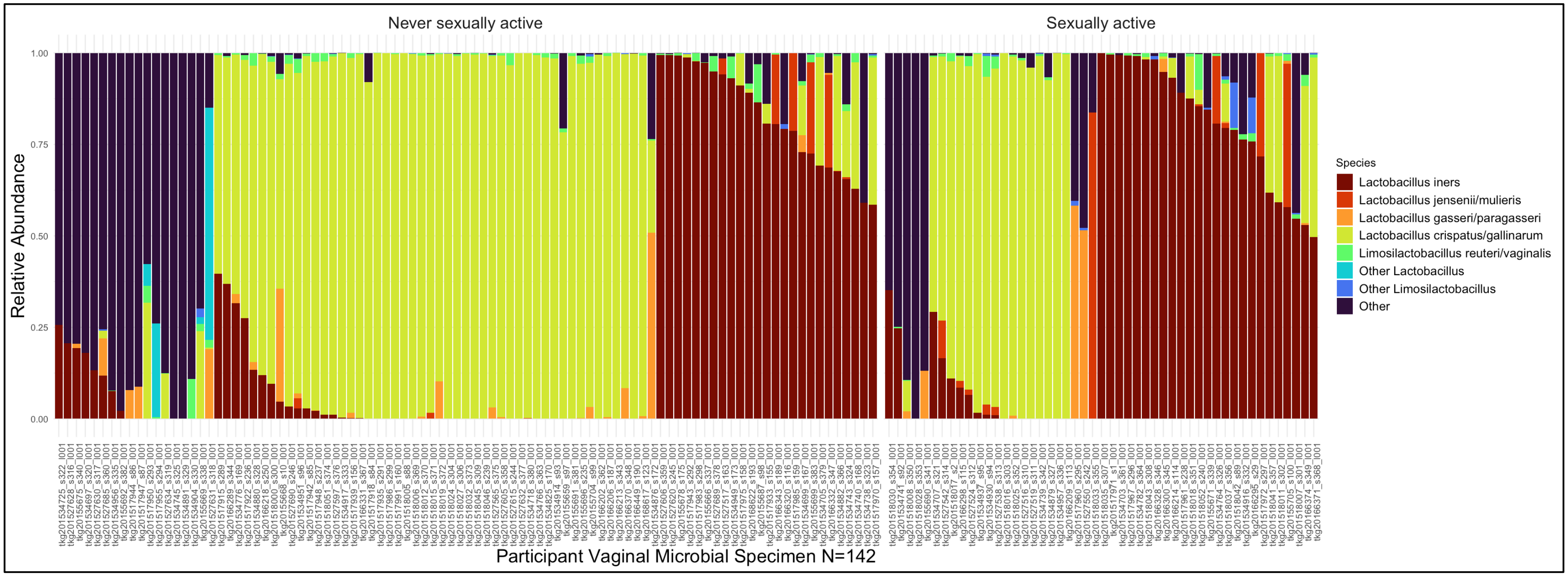

Figure 2. Stacked taxonomic bar plot of relative abundance of key *Lactobacillus* and *Limosilactobacillus* species in the vaginal microbiome of 142 Kenyan adolescent girls and young women. Each bar represents the microbial composition of one specimen. Taxa are grouped by genus/species and participants are grouped by sexual activity.

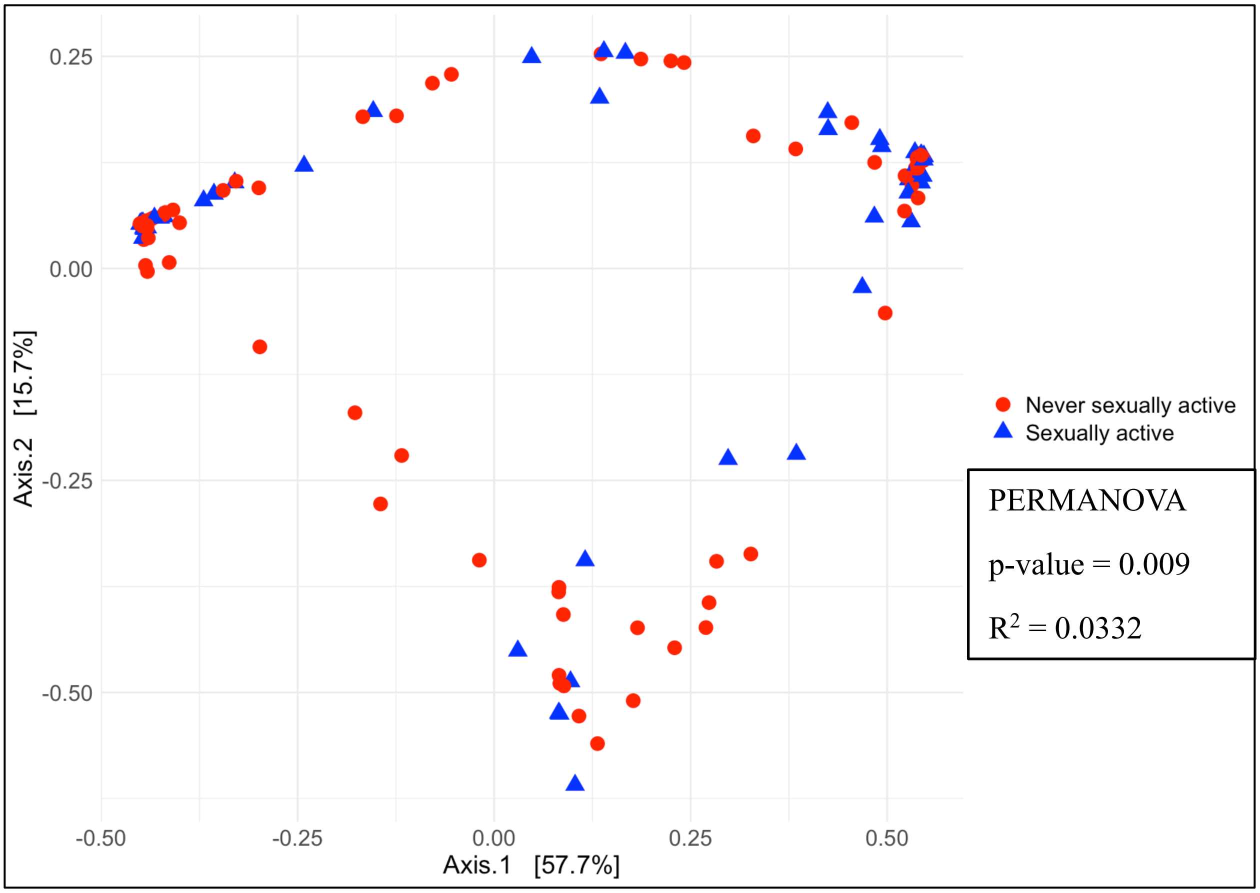

Figure 3. Bray-Curtis Dissimilarity Principal coordinate analysis (PCoA) plot and PERMANOVA of 142 vaginal microbial specimen of Kenyan adolescent girls and young women. Each point represents one specimen, microbial composition is visualized in two dimensions, Axis 1 and 2 explain 57.7% and 15.7% of the variation, respectively. Red circles represent specimen from participants with no previous sexual activity and blue triangles represent specimen from sexually active participants. Higher value on Axis 1 represents more L. iners, L. jensenii/mulieris and less L. crispatus/gallinarum. Higher value on Axis 2 represents more L. iners, L. jensenii/mulieris and more L. crispatus/gallinarum.

Table 2 (a). Unadjusted Prevalence and Poisson Regression Models of Relationship between Detection of *Lactobacillus* and *Limosilactobacillus* species and having engaged in sexual activity compared to not having engaged in sexual activity among Kenyan adolescent girls and young women.

Detection	**Unweighted**				**Weighted**			
Lactobacillus spp.	**Never sexually active N=93 (%)**	**Sexually active N=49 (%)**	**PR (95% CI)**	***p***	**Never sexually active N=599.2 (%)**	**Sexually active N=308.9 (%)**	**PR (95% CI)**	***p***
Lactobacillus crispatus/gallinarum	66 (71.0)	27 (55.1)	0.78 (0.58–1.04)	0.087	469.6 (78.4)	209.6 (67.9)	0.87 (0.60–1.25)	0.439
Lactobacillus gasseri/paragasseri	23 (24.7)	9 (18.4)	0.74 (0.37–1.50)	0.406	101.0 (16.9)	58.8 (19.0)	1.13 (0.34–3.74)	0.843
Lactobacillus iners	50 (53.8)	35 (71.4)	**1.33** **(1.02–1.73)**	**0.034**	316.1 (52.8)	186.9 (60.5)	1.15 (0.68–1.92)	0.603
Lactobacillus jensenii/mulieris	10 (10.8)	11 (22.4)	2.09 (0.94–4.62)	0.070	74.6 (12.5)	96.31 (31.2)	2.50 (0.71–8.88)	0.158
Limosilactobacillus spp.								
Limosilactobacillus coleohominis	18 (19.4)	14 (28.6)	1.48 (0.80–2.73)	0.216	82.1 (13.7)	71.4 (23.1)	1.69 (0.62–4.59)	0.307
Limosilactobacillus reuteri/vaginalis	66 (71.0)	34 (69.4)	0.98 (0.78–1.23)	0.848	446.8 (75.1)	221.7 (71.8)	0.96 (0.67–1.37)	0.805

Table 2 (b). Poisson Regression Models of Relationship between Detection of *Lactobacillus* and *Limosilactobacillus* species and having engaged in sexual activity relative to not having engaged in sexual activity among Kenyan adolescent girls and young women. Adjusted for Age, Rural/Urban Residence, and Housing Condition.

Detection	**Sexually active compared to Never sexually active**			
	Unweighted Adjusted PR (95% CI)	***p***	**Weighted Adjusted PR (95% CI)**	***p***
Lactobacillus spp.				
Lactobacillus crispatus/gallinarum	0.77 (0.57–1.05)	0.103	0.77 (0.54-1.10)	0.155
Lactobacillus gasseri/paragasseri	0.75 (0.34–1.66)	0.480	0.74 (0.28-1.93)	0.538
Lactobacillus iners	**1.36** **(1.02–1.83)**	**0.039**	1.49 (0.87-2.55)	0.148
Lactobacillus jensenii/mulieris	**2.44** **(1.12–5.29)**	**0.024**	**3.09** **(1.27-7.52)**	**0.014**
Limosilactobacillus spp.				
Limosilactobacillus coleohominis	1.58 (0.80–3.11)	0.190	**2.86** **(1.03-7.93)**	**0.046**
Limosilactobacillus reuteri/vaginalis	1.01 (0.78–1.31)	0.940	1.07 (0.76-1.51)	0.704

Table 3. Compositional LASSO analysis describing the odds group membership per increase in relative abundance of vaginal bacterial species.

	Sexually active compared to Never sexually active
	Odds Ratio[**]
Species[*]	
Lactobacillus crispatus/gallinarum	0.978
Lactobacillus iners	1.022

*Other species in model included: *L. gasseri/paragasseri, L. jensenii/mulieris, L. delbrueckii, L. equicursoris, L. johnsonii, L. kitasatonis, L. taiwanensis, L. coleohominis, L. fermentum, L. reuteri/vaginalis;* all other species categorized into group "Other" and included in the model.

**Odds ratios (OR) per log2 increase in relative abundance of species using compositional LASSO logistic regression. Sexually active (Y=1) versus Never sexually active (Y=0). Zeros in the count table were replaced with pseudo count of 0.5. Compositional LASSO model was fit using the zeroSum package version 2.0.7 (https://github.com/rehbergT/zeroSum) in R version 4.3.3, using the lambda within one standard error (SE) of the minimum mean-square error (MSE).

Table 4 (a). Demographic, socioeconomic, and health correlates of detection of *Limosilactobacillus* species using Poisson Regression. All variables shown were included simultaneously in the model. Prevalence ratios (PR) and 95% confidence intervals (CI) are reported for both weighted and unweighted models. Estimates are adjusted for all listed covariates.

Characteristics	Detection of *Limosilactobacillus* species			
	Unweighted Adjusted PR (95% CI)	***p***	**Weighted Adjusted PR (95% CI)**	***p***
Age (years)	0.97 (0.88 – 1.06)	0.443	0.91 (0.80 – 1.04)	0.182
Education (years)	1.04 (0.97 – 1.12)	0.187	0.97 (0.90 – 1.04)	0.372
Residence				
Rural	Ref.			
Urban	1.05 (0.86 – 1.28)	0.590	0.90 (0.70 – 1.17)	0.449
Housing Condition †				
High	Ref.			
Medium	1.11 (0.85 – 1.44)	0.412	0.68 (0.42 – 1.10)	0.118
Low	0.98 (0.80 – 1.20)	0.829	**0.73** **(0.53 – 1.00)**	**0.050**
Ever Sexually Active				
No	Ref.			
Yes	1.01 (0.83 – 1.24)	0.893	1.11 (0.84 – 1.48)	0.462
Nugent Score				
0-3	Ref.			
4-6	0.57 (0.25 – 1.30)	0.117	0.58 (0.24 – 1.44)	0.243
7-10 (BV)	0.45 (0.16 – 1.27)	0.078	**0.19** **(0.05 – 0.81)**	**0.026**

Table 4 (b). Associations between detection of *Limosilactobacillus* species and other vaginal bacterial species, using Poisson regression. All bacterial species listed were included in the same model. Prevalence ratios (PR) and 95% confidence intervals (CI) reported for both unweighted and weighted models. Estimates are adjusted for all species shown.

Species	**Detection of *Limosilactobacillus* species**			
	Unweighted Adjusted PR (95% CI)	***p***	**Weighted Adjusted PR (95% CI)**	***p***
Gardnerella	0.96 (0.79 – 1.16)	0.643	0.92 (0.66 – 1.27)	0.599
Lactobacillus crispatus/gallinarum	**1.51** **(1.17 – 1.95)**	**0.002**	1.50 (0.96 – 2.35)	0.077
Lactobacillus iners	**1.22** **(1.02 – 1.47)**	**0.029**	1.34 (0.99 – 1.81)	0.060
Lactobacillus gasseri/paragasseri	**1.20** **(1.02 – 1.42)**	**0.029**	1.29 (0.99 – 1.67)	0.057
Lactobacillus jensenii/mulieris	0.89 (0.68 – 1.18)	0.424	1.02 (0.79 – 1.32)	0.873

Citations

1. World Health Organization. Implementing the Global Health Sector Strategies on HIV, Viral Hepatitis and Sexually Transmitted Infections, 2022–2030: Report on Progress and Gaps 2024. Geneva: World Health Organization; 2024. Licence: CC BY-NC-SA 3.0 IGO.

2. Eastment MC, McClelland RS. Vaginal microbiota and susceptibility to HIV. AIDS. 2018;32(6):687-698. doi:10.1097/QAD.0000000000001768

3. Brotman RM. Vaginal microbiome and sexually transmitted infections: an epidemiologic perspective. J Clin Invest. 2011;121(12):4610-4617. doi:10.1172/JCI57172

4. Abou Chacra L, Fenollar F, Diop K. Bacterial Vaginosis: What Do We Currently Know? Front Cell Infect Microbiol. 2021;11:672429. doi:10.3389/fcimb.2021.672429

5. Jespers V, Hardy L, Buyze J, Loos J, Buvé A, Crucitti T. Association of Sexual Debut in Adolescents With Microbiota and Inflammatory Markers. Obstetrics & Gynecology. 2016;128(1):22-31. doi:10.1097/AOG.0000000000001468

6. Lebeer S, Ahannach S, Gehrmann T, et al. A citizen-science-enabled catalogue of the vaginal microbiome and associated factors. Nat Microbiol. 2023;8(11):2183-2195. doi:10.1038/s41564-023-01500-0

7. Ratten LK, Plummer EL, Murray GL, et al. Sex is associated with the persistence of non-optimal vaginal microbiota following treatment for bacterial vaginosis: a prospective cohort study. BJOG. 2021;128(4):756-767. doi:10.1111/1471-0528.16430

8. for the Vaginal Biomarkers Study Group, Jespers V, Van De Wijgert J, et al. The significance of Lactobacillus crispatus and L. vaginalis for vaginal health and the negative effect of recent sex: a cross-sectional descriptive study across groups of African women. BMC Infect Dis. 2015;15(1):115. doi:10.1186/s12879-015-0825-z

9. Cherpes TL, Hillier SL, Meyn LA, Busch JL, Krohn MA. A Delicate Balance: Risk Factors for Acquisition of Bacterial Vaginosis Include Sexual Activity, Absence of Hydrogen Peroxide-Producing Lactobacilli, Black Race, and Positive Herpes Simplex Virus Type 2 Serology. Sexually Transmitted Diseases. 2008;35(1):78-83. doi:10.1097/OLQ.0b013e318156a5d0

10. Lewis FMT, Bernstein KT, Aral SO. Vaginal Microbiome and Its Relationship to Behavior, Sexual Health, and Sexually Transmitted Diseases. Obstet Gynecol. 2017;129(4):643-654. doi:10.1097/AOG.0000000000001932

11. Bayigga L, Kateete DP, Anderson DJ, Sekikubo M, Nakanjako D. Diversity of vaginal microbiota in sub-Saharan Africa and its effects on HIV transmission and prevention. American Journal of Obstetrics and Gynecology. 2019;220(2):155-166. doi:10.1016/j.ajog.2018.10.014

12. McClelland RS, Lingappa JR, Srinivasan S, et al. Evaluation of the association between the concentrations of key vaginal bacteria and the increased risk of HIV acquisition in African women from five cohorts: a nested case-control study. Lancet Infect Dis. 2018;18(5):554-564. doi:10.1016/S1473-3099(18)30058-6

13. Borgdorff H, Tsivtsivadze 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(9):1781-1793. doi:10.1038/ismej.2014.26

14. Happel AU, Balle C, Havyarimana E, et al. Cervicovaginal Human Papillomavirus Genomes, Microbiota Composition and Cytokine Concentrations in South African Adolescents. Viruses. 2023;15(3):758. doi:10.3390/v15030758

15. Rahman N, Mian MF, Hayes CL, Nazli A, Kaushic C. G. vaginalis increases HSV-2 infection by decreasing vaginal barrier integrity and increasing inflammation in vivo. Front Immunol. 2024;15:1487726. doi:10.3389/fimmu.2024.1487726

16. Jakobsson T, Forsum U. Lactobacillus iners: a marker of changes in the vaginal flora? J Clin Microbiol. 2007;45(9):3145. doi:10.1128/JCM.00558-07

17. Jespers V, Kyongo J, Joseph S, et al. A longitudinal analysis of the vaginal microbiota and vaginal immune mediators in women from sub-Saharan Africa. Sci Rep. 2017;7(1):11974. doi:10.1038/s41598-017-12198-6

18. Zheng N, Guo R, Wang J, Zhou W, Ling Z. Contribution of Lactobacillus iners to Vaginal Health and Diseases: A Systematic Review. Front Cell Infect Microbiol. 2021;11:792787. doi:10.3389/fcimb.2021.792787

19. Vodstrcil LA, Twin J, Garland SM, et al. The influence of sexual activity on the vaginal microbiota and Gardnerella vaginalis clade diversity in young women. PLoS One. 2017;12(2):e0171856. doi:10.1371/journal.pone.0171856

20. Zheng J, Wittouck S, Salvetti E, et al. A taxonomic note on the genus Lactobacillus: Description of 23 novel genera, emended description of the genus Lactobacillus Beijerinck 1901, and union of Lactobacillaceae and Leuconostocaceae. International Journal of Systematic and Evolutionary Microbiology. 2020;70(4):2782-2858. doi:10.1099/ijsem.0.004107

21. Yuh T, Micheni M, Selke S, et al. Sexually Transmitted Infections Among Kenyan Adolescent Girls and Young Women With Limited Sexual Experience. Front Public Health. 2020;8:303. doi:10.3389/fpubh.2020.00303

22. Alishum A. DADA2 formatted 16S rRNA gene sequences for both bacteria & archaea. Published online January 16, 2019. doi:10.5281/ZENODO.3188334

23. Callahan B. RDP taxonomic training data formatted for DADA2 (RDP trainset 18/release 11.5). Published online December 7, 2020. doi:10.5281/ZENODO.4310151

24. McLaren MR, Callahan BJ. Silva 138.1 prokaryotic SSU taxonomic training data formatted for DADA2. Published online March 7, 2021. doi:10.5281/ZENODO.4587955

25. 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. Fredricks DN, ed. PLoS ONE. 2015;10(2):e0116894. doi:10.1371/journal.pone.0116894

26. Francis SC, Crucitti T, Smekens T, et al. The Vaginal Microbiota Among Adolescent Girls in Tanzania Around the Time of Sexual Debut. Front Cell Infect Microbiol. 2020;10:305. doi:10.3389/fcimb.2020.00305

27. Novak J, Ravel J, Ma B, et al. Characteristics associated with Lactobacillus iners-dominated vaginal microbiota. Sex Transm Infect. 2022;98(5):353-359. doi:10.1136/sextrans-2020-054824

28. Mehta SD, Zulaika G, Otieno FO, et al. High Prevalence of Lactobacillus crispatus Dominated Vaginal Microbiome Among Kenyan Secondary School Girls: Negative Effects of Poor Quality Menstrual Hygiene Management and Sexual Activity. Front Cell Infect Microbiol. 2021;11:716537. doi:10.3389/fcimb.2021.716537

29. Putonti C, Shapiro JW, Ene A, Tsibere O, Wolfe AJ. Comparative Genomic Study of Lactobacillus jensenii and the Newly Defined Lactobacillus mulieris Species Identifies Species-Specific Functionality. Rao K, ed. mSphere. 2020;5(4):10.1128/msphere.00560-20. doi:10.1128/msphere.00560-20

30. Chee WJY, Chew SY, Than LTL. Vaginal microbiota and the potential of Lactobacillus derivatives in maintaining vaginal health. Microb Cell Fact. 2020;19(1):203. doi:10.1186/s12934-020-01464-4

31. Witkin SS, Mendes-Soares H, Linhares IM, Jayaram A, Ledger WJ, Forney LJ. Influence of Vaginal Bacteria and D - and L -Lactic Acid Isomers on Vaginal Extracellular Matrix Metalloproteinase Inducer: Implications for Protection against Upper Genital Tract Infections. Blaser MJ, ed. mBio. 2013;4(4):e00460-13. doi:10.1128/mBio.00460-13

32. Sabo MC, Lokken EM, Srinivasan S, et al. Changes in Vaginal Bacteria and Inflammatory Mediators From Periconception Through the Early Postpartum Period in a Cohort of Kenyan Women Without HIV. The Journal of Infectious Diseases. 2023;228(4):487-499. doi:10.1093/infdis/jiad168

33. Shazadi K, Arshad N. Evaluation of inhibitory and probiotic properties of lactic acid bacteria isolated from vaginal microflora. Folia Microbiol (Praha). 2022;67(3):427-445. doi:10.1007/s12223-021-00942-5

34. Song SD, Acharya KD, Zhu JE, et al. Daily Vaginal Microbiota Fluctuations Associated with Natural Hormonal Cycle, Contraceptives, Diet, and Exercise. mSphere. 2020;5(4):e00593-20. doi:10.1128/mSphere.00593-20

35. Srinivasan S, Liu C, Mitchell CM, et al. Temporal variability of human vaginal bacteria and relationship with bacterial vaginosis. PLoS One. 2010;5(4):e10197. doi:10.1371/journal.pone.0010197