Factors associated with the composition and diversity of the cervical microbiota of reproductive-age Black South African women: a retrospective cross-sectional study

Harris Onywera[1,2], Anna-Lise Williamson[1,2,3], Zizipho Z.A. Mbulawa[1,2,3,4], David Coetzee[5] and Tracy L. Meiring[1,2]

[1] Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town, Western Cape, South Africa

[2] Division of Medical Virology, Department of Pathology, Faculty of Health Sciences, University of Cape Town, Cape Town, Western Cape, South Africa

[3] SAMRC Gynaecological Cancer Research Centre, University of Cape Town, Cape Town, Western Cape, South Africa

[4] Center for HIV & STIs, National Institute for Communicable Diseases, National Health Laboratory Service, University of Cape Town, Johannesburg, Gauteng, South Africa

[5] Center for Infectious Disease Epidemiology and Research, School of Public Health and Family Medicine, University of Cape Town, Cape Town, Western Cape, South Africa

Corresponding author
Harris Onywera,
do79harris@gmail.com

ABSTRACT

Background. *Lactobacillus* spp. are common bacteria in the cervical and vaginal microbiota (CVM) and are thought to represent a "healthy" cervicovaginal state. Several studies have found an independent association between ethnicity/race and cervical and vaginal microbiota (CVM) composition. Women of sub-Saharan African descent appear to be significantly more likely to have non-*Lactobacillus*-dominated CVM compared to women of European descent. The factors contributing to these differences remain to be fully elucidated. The CVM of Black South African women and factors influencing their CVM remain understudied. In this study, we characterized the cervical microbiota of reproductive-age South African women and assessed the associations of these microbiota with participants' metadata.

Methods. The cervical microbiota from cervical DNA of 62 reproductive-age women were profiled by Ion Torrent sequencing the V4 hypervariable region of the bacterial 16S ribosomal RNA (rRNA) gene and analyzed with the Quantitative Insights Into Microbial Ecology (QIIME), UPARSE, and metagenomeSeq tools. Associations between cervical microbiota and participants' metadata were assessed using GraphPad Prism, R packages and an in-house script.

Results. The cervical microbiota clustered into three distinct community state types (CSTs): *Lactobacillus iners*-dominated cervical microbiota (CST I (38.7%, 24/62)), unclassified *Lactobacillus*-dominated cervical microbiota (CST II (4.8%, 3/62)), and diverse cervical microbiota (CST III (56.5%, 35/62)) with an array of heterogeneous bacteria, predominantly the bacterial vaginosis (BV)-associated *Gardnerella*, *Prevotella*, *Sneathia*, and *Shuttleworthia*. CST III was associated with BV ($p = 0.001$). Women in CST I were more likely to be on hormonal contraception, especially progestin-based, compared to women in CST III (odds ratio: 5.2 (95% CI [1.6–17.2]); $p = 0.005$).

Women on hormonal contraception had a significantly lower alpha (Shannon indices: 0.9 (0.2–1.9) versus 2.3 (0.6–2.3); $p = 0.025$) and beta (permutational multivariate analysis of variance (PERMANOVA) pseudo-F statistic = 4.31, $p = 0.019$) diversity compared to non-users. There was no significant difference in the alpha (Shannon indices: 1.0 (0.3–2.2) versus 1.9 (0.3–2.2); $p = 0.483$) and beta (PERMANOVA pseudo-F statistic = 0.89, $p = 0.373$) diversity in women with versus without human papillomavirus infection.

Conclusions. The majority of Black women in our study had non-*Lactobacillus*-dominated cervical microbiota. Additional studies are needed to examine whether such microbiota represent abnormal, intermediate or variant states of health. Lastly, the association of hormonal contraception with *L. iners* dominance requires further in-depth research to confirm this association, determine its biological mechanism and whether it has a beneficial effect on the cervicovaginal health.

INTRODUCTION

It is generally acknowledged that *Lactobacillus* spp., particularly *Lactobacillus crispatus*, *Lactobacillus gasseri*, *Lactobacillus iners*, and *Lactobacillus jensenii*, are common bacteria in the cervical and vaginal microbiota (CVM) (*Huang et al., 2015*; *Ravel et al., 2011*; *Smith et al., 2012*) and are regarded as biomarkers of health (*Huang et al., 2015*; *Petrova et al., 2015*; *Ravel et al., 2011*). Lactobacilli are Gram-positive bacteria (*Hillier, 1993*) that are thought to have a protective role in preventing genital disease by restricting the growth of non-indigenous organisms, including sexually transmitted infections (STIs) (*Borgdorff et al., 2014*; *Castro et al., 2013*; *Petrova et al., 2015*). *L. gasseri*, for instance, has been associated with rapid clearance of human papillomavirus (HPV) infection (*Brotman et al., 2014b*). Loss of lactobacilli concomitantly with overgrowth of anaerobic and microaerophilic bacteria, which can be Gram-positive, Gram-negative, and/or Gram-variable bacteria, results in bacterial vaginosis (BV) (*Hillier, 1993*; *Sobel, 2000*), the most common form of vaginal disorder among reproductive-age women (*Kenyon, Colebunders & Crucitti, 2013*). BV has been associated with increased risk of acquisition of a range of STIs such as HPV (*Nardis, Mosca & Mastromarino, 2013*; *Watts et al., 2005*) that is associated with cervical neoplasia (*Koshiol et al., 2008*) and cervical cancer (*Bosch et al., 2002*). South African women have a high prevalence of BV (31–63%) (*Abbai, Reddy & Ramjee, 2016*; *Kenyon, Colebunders & Crucitti, 2013*; *Lennard et al., 2018*; *Onywera et al., 2019*) and HPV (16–75%) (*Bruni et al., 2019*) yet there is paucity of knowledge on the CVM of Black South African women as the existing CVM studies have mostly focussed on women of White, Asian, Hispanic, and African American background.

It is currently known that the composition of the CVM is not only affected by factors ranging from smoking, sexual behaviour, menstrual cycle, and hormonal contraception practices to pregnancy (*Borgdorff et al., 2017*; *Brotman et al., 2014a*; *Gajer et al., 2012*;

Romero et al., 2014; *Vodstrcil et al., 2017*; *Wessels et al., 2017*), but ethnic/racial background as well (*Borgdorff et al., 2017*; *Fettweis et al., 2014*; *Ravel et al., 2011*; *Vodstrcil et al., 2017*; *Zhou et al., 2007*). A large extensive study on 396 asymptomatic reproductive-age multi-ethnic North American women observed different prevalences of *Lactobacillus*-dominated vaginal microbiota among White (89.7%), Asian (80.2%), Hispanic (61.9%), and Black (59.6%) women (*Ravel et al., 2011*). Most studies have reported that 23–64% of reproductive-age Black women have diverse CVM with low numbers of *Lactobacillus* sp. that are credited for sustaining health (*Anahtar et al., 2015*; *Borgdorff et al., 2014*; *Dareng et al., 2016*; *Fettweis et al., 2014*; *Gautam et al., 2015*; *Lennard et al., 2018*; *Onywera et al., 2019*; *Ravel et al., 2011*; *Wessels et al., 2017*; *Zhou et al., 2007*). Among women of non-Black ethnicity, this prevalence seldom reaches 40% (*Fettweis et al., 2014*; *Ravel et al., 2011*; *Zhou et al., 2007*). Controversy abounds whether there are actual differences in the CVM composition among ethnic/racial groups or the higher prevalences of diverse non-*Lactobacillus*-dominated CVM in some populations (including Hispanic or Black) may reflect higher rates of asymptomatic BV or STIs (*Beamer et al., 2017*).

We now know that among the Black South African women with *Lactobacillus*-dominated CVM, a high proportion of them (59–83%) have *L. iners*-dominated CVM (*Anahtar et al., 2015*; *Balle et al., 2018*; *Lennard et al., 2018*; *Onywera et al., 2019*). This seems contrary to the observation in White women (<45%), where *L. crispatus* is the most predominant *Lactobacillus* spp. (*Fettweis et al., 2014*; *Ravel et al., 2011*; *Zhou et al., 2007*). Of the *Lactobacillus* spp. considered as biomarkers of a healthy cervicovaginal tract (*Petrova et al., 2015*; *Ravel et al., 2011*), *L. iners* appears to be the least stable (*Gajer et al., 2012*; *Petrova et al., 2017*) and least protective against BV and STIs (*Brotman et al., 2014b*; *Petrova et al., 2017*; *Van Houdt et al., 2018*; *Verstraelen et al., 2009*). More studies are needed to understand the CVM and factors influencing them. Therefore, owing to the reports on ethnic/racial variations in the CVM and dearth of knowledge about the CVM of Black South African women, we aimed to investigate the baseline structure of cervical microbiota of reproductive-age Black South African women and determine their (microbiota) associations with the participants' demographic, sociobehavioural, and clinical information.

MATERIALS AND METHODS

Ethics statement
This study was approved by the Human Research Ethics Committee of the University of Cape Town, South Africa (references 258/2006 and 580/2014). All participants provided written informed consent to participate in the study and use of their stored samples for future studies.

Study population and study design
This was a retrospective cross-sectional study based on data and baseline cervical DNA samples from the HPV Couples Cohort Study (*Mbulawa et al., 2009*) that examined the transmission of genital HPV among Black heterosexual couples in Gugulethu, Cape Town,

South Africa. Details of enrollment, recruitment and sample collection for the HPV transmission study have been described previously (*Mbulawa et al., 2009*).

In brief, speculum examination was performed and excess mucus around the cervical area was cleared using a filamented swab. This was followed by collection of two cytobrush samples from the cervix. The first sample was for Papanicolaou (Pap) smear or T cell assay. It was collected by inserting the cytobrush into the mouth of the cervix and rotating the cytobrush at 360° thrice. For the Pap smear, the sample was smeared immediately onto the frosted glass slide, quickly fixed using Cytofix spray and then stained with Pap stain. BV was identified on Pap smears by using the Bethesda criteria for reporting cervical/vaginal cytologic diagnoses (*Kurman & Solomon, 1994*). Smears showing clue cells with coccobacilli (mostly *Gardnerella vaginalis*) and/or any shifts in bacterial flora suggestive of BV (noticeable absence of lactobacilli) on wet microscopy were considered as having findings suggestive of BV. All smears read by cytotechnologists were reviewed. The second sample from the cervix was for HPV genotyping and herpes simplex virus (HSV) testing (and subsequent analyses such as characterization of microbiome). It was collected by inserting a second cytobrush (Digene cervical sampler) into the cervix and rotating it thrice (360°) inside the mouth of the cervix. This sample was then stored in Digene specimen transport medium (Digene Corporation, Gaithersburg, MD, USA) at −80 °C until nucleic acid extraction.

Nucleic acids were extracted from the cervical samples as previously described (2009) (*Mbulawa et al., 2009*) using the MagNA Pure Compact System and the MagNA Pure Compact Nucleic Acid Isolation kit (Roche Molecular Diagnostics, Mannheim, Germany). HPV typing was also performed as previously documented (*Mbulawa et al., 2009*) using the Roche Linear Array HPV genotyping test (Roche Molecular Diagnostics, Mannheim, Germany) that detects 37 HPV genotypes. These include 12 oncogenic high-risk, 8 probable oncogenic high-risk, and 17 non-oncogenic low-risk HPV types as listed elsewhere (*Mbulawa et al., 2009*). Only samples with positive human beta (β)-globin (a housekeeping gene) hybridization results (a measure of sample adequacy) were included in this study. Roche Linear Array HPV genotyping test measures sample adequacy by relying on two endogenous β-globin positive controls (high and low) run concurrently with samples. The primers targeting β-globin are different from those that target the HPV genome (polymorphic L1 region). A valid genotyping result (negative or positive for HPV genotype) is one where both the β-globin probe lines are positive. A result (negative or positive for at least one HPV genotype) is considered invalid if the sample is negative for one or both β-globin control(s). This is suggestive of inadequate cellular material, poor storage and processing (extraction), presence of PCR inhibitors and/or completion with a high titer HPV target.

To be eligible for the present study, only cervical DNA specimens from human immunodeficiency virus (HIV)-seronegative women aged 18-44 years were considered. These samples had to have information on the HPV status and sufficient volume for microbiota analysis (≥15.0 µl of the extracted DNA). Exclusion criteria included being a woman aged <18 or >44 years, self-reported menstruation or pregnancy at the time of sampling, and HIV-seropositivity. Participants' metadata including demographics, sexual

history, smoking, contraceptive use, and clinical characteristics were abstracted from the HPV Couples Cohort Study (*Mbulawa et al., 2009*).

Bacterial V4 hypervariable region (16S ribosomal rRNA) library preparation and sequencing

The hypervariable V4 region of the 16S ribosomal rRNA (rRNA) gene was amplified using the universal polymerase chain reaction (PCR) primers 515f (5′-GTGCCAGCMGCCGCG GTAA-3′) and 806r (5′-GGACTACHVGGGTWTCTAAT-3′) (*Caporaso et al., 2011*). Each PCR contained 1x Ex Taq buffer (Takara Bio Inc., Japan), 0.025 U Ex Taq polymerase, 0.8 mM deoxynucleotide triphosphate (dNTP) mixture, 0.56 mg/ml bovine serum albumin (BSA), 400 nM each primer and 100 ng template. Each sample (and no template PCR control, i.e., nuclease free water) was amplified in 3 replicate reactions. PCR conditions were 98 °C for 2 min, followed by 30 cycles of 98 °C for 20 s, 50 °C for 30 s and 72 °C for 45 s, and a final elongation step at 72 °C for 10 min. The triplicate samples were pooled and purified using the Agencourt AMPure XP system (Beckman Coulter, Germany) according to the manufacturer's instructions. Amplicon sizes were confirmed by electrophoresis on 1.5% Tris Borate EDTA (TBE) agarose gels and imaging with the ultraviolet transilluminator (UVT) GelDoc-It[TM] system. The amplicons were quantified using the Quant-iT[®] PicoGreen dsDNA assay (Thermo Fisher Scientific, USA) with FLUOstar OPTIMA Multi-Mode Micro Plate Reader (BMG Labtech, Germany). Sequencing libraries were prepared using the KAPA Library Preparation kit and barcoded Adaptor kits for Ion Torrent[TM] platforms (KAPA Biosystems, Wilmington, MA, USA). Purified barcoded amplicons were pooled (24 samples per pool) in equimolar amounts and the final library sizes and concentrations assessed on a Bioanalyzer High Sensitivity Chip (Agilent Technologies, Santa Clara, CA, USA). Sequencing was performed on the Ion Torrent Personal Genome Machine (PGM) (Life Technologies, Beverly, MA, USA) at the Central Analytical Facilities (CAF) at Stellenbosch University (Stellenbosch, South Africa).

Bacterial V4 hypervariable region (16S rRNA) data analysis using bioinformatics tools

The qualities of the raw sequenced reads were visually inspected using FastQC v0.11.2 (*Andrews, 2010*). Quantitative Insights Into Microbial Ecology (QIIME) v1.8.0 (*Caporaso et al., 2010b*) with imported UPARSE (usearch7.0.1090) (*Edgar, 2013*), was used to analyze and interpret the nucleotide sequence data from the cervical microbiota. In the initial sequence pass, reads were quality-filtered and demultiplexed in QIIME using the user-defined parameters in Table S1. Briefly, reads with lengths outside the 200-400-bp range, with a quality score of <25 (sliding window 50) and without barcodes or with any mismatches in the barcode sequences were discarded. A second quality filter was performed in UPARSE with user-defined parameters (Table S1). Sequences were dereplicated followed by abundance sorting and discarding singletons. Operational taxonomic unit (OTU) clustering was performed by UPARSE-OTU method that uses a greedy clustering algorithm, with binning of reads with 97% pairwise identity. This step was performed simultaneously with representative sequence picking and *de novo* chimera filtering. Representative sequences from each unique OTU cluster were picked using

abundance algorithm. Additional chimeras were removed by UCHIME algorithm (*Edgar et al., 2011*). Taxonomy was assigned using the Ribosomal Database Project (RDP) Naïve Bayesian Classifier (*Wang et al., 2007*), with the Greengenes database (gg13_8 Release) (*De Santis et al., 2006*). Phylogeny was inferred by aligning representative sequences to Greengenes core set using Python Nearest Alignment Space Termination (PyNAST) (*Caporaso et al., 2010a*). A phylogenetic tree was built using FastTree. Other parameters used for our analyses are defined in Table S1.

Diversity, rarefaction, and sample ordinations were computed in QIIME. Multiple rarefactions at different sequencing depths were performed, and rarefactions (collector's) curves plotted to evaluate the completeness of the sampling efforts. Alpha diversity was computed by chao1, observed_species, Shannon, Simpson, and PD_whole_tree metrics. Beta diversity was estimated using weighted and unweighted UniFrac distances, and Bray-Curtis dissimilarity metric. The strength and statistical significance of sample clustering (beta diversity) was computed using permutational multivariate analysis of variance (PERMANOVA) (*Anderson, 2001*), with 999 permutations. Other biodiversity metrics including Dominance and Shannon Equitability indices were calculated using an in-house script in RStudio v1.1.447 (*RStudio Team, 2016*). An all-by-all pairwise distance matrix of UniFrac distances were generated and used to hierarchically cluster and ordinate samples. The ordinations were performed using Principal Coordinate Analysis (PCoA).

Identification and comparison of community state types

Hierarchical clustering and heatmap generation were performed in RStudio v1.1.447 (*RStudio Team, 2016*). Hierarchical clustering on Bray-Curtis dissimilarity (Vegan package v2.5-5 (*Philip, 2003*)) was done using the average neighbour algorithm. The heatmap was generated using the heatmap3 v1.1.6 package (*Zhao et al., 2014*).

Correlational analyses of 60 cervical bacterial OTUs

To assess bacterial positive and negative relationships, Spearman's correlation of $\log_2$-transfomed counts of all OTU pairs were tested on metagenomeSeq v1.12.1 (*Paulson et al., 2014*; *Paulson et al., 2016*) in RStudio v1.1.447 (*RStudio Team, 2016*), and the topmost OTUs ($n = 60$, 28.0%) with the greatest variance displayed on a correlogram.

Statistical methods

Statistical analyses were carried out using GraphPad Prism v6.01 (San Diego, CA, USA). Mann–Whitney unpaired nonparametric and Chi-square/Fisher's exact tests (with two-tailed *p*-value) were used to examine the association of continuous and categorical variables with CSTs, as appropriate. The alpha diversity metrics of the three CSTs were compared by Kruskal-Wallis test. Two-group comparison between the alpha diversity of CST, HPV, and BV groups was computed by Mann–Whitney unpaired nonparametric test.

RESULTS

Study cohort baseline characteristics

The demographic, sexual and smoking, behavioural and clinical information of the 62 heterosexual Black South African women included in this study are summarized in Table 1.

All the women were sexually active, with the majority (72%) of hormonal contraceptives users being on Depo-Provera. HPV and high-risk HPV infections were detected in 37.1% and 29.0% of the women, correspondingly. Thirteen of the women (22.0%) had abnormal cervical cytology. A few women had experienced vaginal discharge (16.1%) and genital ulceration (3.2%) in the last six months. Twenty two women (35.5%) had findings suggestive of BV. A majority of the women (79.0%) had never smoked cigarettes.

Taxonomic composition of the cervical microbiota

A total of 1,392,562 high-quality non-spurious sequencing reads from 62 samples were included in the final analysis with a median of 16,453 reads per sample (range: 5,343–235,897 per sample). The Ion Torrent PGM raw sequence data and metadata have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject PRJNA472137 (SRP148486) with accession numbers (SRX4103412–SRX4103416).

A heatmap of the relative abundances of the predominant genera (>0.5% relative abundance) and their respective phyla identified in the cervical microbiota of the 62 women is shown in Fig. S1. Above the 0.5% threshold, only 9 phyla with 43 genera were observed. *Lactobacillus* (46.0% mean relative abundance and 100.0% prevalence), *Gardnerella* (19.5% and 93.5% (58/62)), *Prevotella* (9.7% and 95.2% (59/62)), and *Sneathia* (9.5% and 87.1% (54/62)) were the most predominant genera in the phyla Firmicutes, Actinobacteria, Bacteriodetes, and Fusobacteria, respectively. Other bacteria such as *Dialister* (1.2%), *Atopobium* (0.5%), and *Clostridium* (0.1%) commonly found in women with BV (*Srinivasan et al., 2012*) were not abundant.

A total of 221 unique OTUs (potential species) were identified in the 62 cervical microbiota and ranged from 16-104 OTUs per sample. The 30 most abundant OTUs represented 97% of all the reads. The most abundant OTU was classified as *L. iners*, representing 40.6% of all the reads. The prevalence and abundance of this OTU together with those of the other most abundant OTUs are shown in Table S2.

Prevotella was the most diverse genus with 26 different OTUs identified. Six and eight different OTUs belonging to the genera *Gardnerella* and *Lactobacillus*, respectively, were detected. Apart from *L. iners*, the other *Lactobacillus* OTUs included *Lactobacillus coleohominis*, *Lactobacillus mucosae*, *Lactobacillus ruminis*, and four unclassified *Lactobacillus* spp. Species names could not be assigned to the four *Lactobacillus* OTUs due to insufficient taxonomic discrimination by the V4 hypervariable region of the 16S rRNA gene. Forty seven (75.8%) of the cervical microbiota had at least two *Lactobacillus* spp., but at unequal abundances.

Characterization of cervical community state types

Hierarchical clustering of the cervical microbiota based on the type and relative abundances of the bacterial taxa identified three distinct community state types, CSTs I–III (Fig. 1). CST I was dominated by *L. iners* and found in 24 women (38.7%). CST II was dominated by an unclassified *Lactobacillus* (*Lactobacillus.4*) and present in only three women (4.8%). CST III was the most common CST occurring in 35 women (56.5%). This CST was characterized by a diverse and complex array of facultative and strictly anaerobic BV-associated bacteria
Table 1 Baseline demographic, sociobehavioural and clinical information of the 62 heterosexual Black South African women.

Characteristic	All participants (N = 62)
Age (years)	34.5 (25.8–39.0)
Age at sexual debut (years)[a]	18.0 (17.0–18.8)
Lifetime number of sexual partners[a]	2.0 (2.0–4.0)
Number of sex acts with study partner in last month[a]	2.0 (2.0–4.0)
Current use of hormonal contraceptives* (% (n/N))	
No	55.4 (31/56)
Yes	44.6 (25/56)
Type of hormonal contraceptives (% (n/N))	
Depo-Provera[b]	72.0 (18/25)
Nonrethisterone enanthate[b]	20.0 (5/25)
Oral pills[c]	8.0 (2/25)
HPV (% (n/N))	
Negative	62.9 (39/62)
Positive	37.1 (23/62)
Multiplicity of HPV infection (% (n/N)) among HPV-infected women	
Single infection	65.2 (15/23)
Multiple infections	34.8 (8/23)
Oncogenicity of HPV infection (% (n/N)) among HPV-infected women	
Low-risk genotypes regardless of high-risk genotype	34.8 (8/23)
High-risk genotypes regardless of low-risk genotype	78.3 (18/23)
Both low- and high-risk genotypes	13.0 (3/23)
High-risk (% (n/N))	
Negative	71.0 (44/62)
Positive	29.0 (18/62)
Cervical cytology (% (n/N))	
Normal	78.0 (46/59)
ASCUS	6.8 (4/59)
LSIL	11.9 (7/59)
HSIL	3.4 (2/59)
Experienced vaginal discharge in last 6 months (% (n/N))	
No	83.9 (52/62)
Yes	16.1 (10/62)
Experienced genital ulceration in last 6 months (% (n/N))	
No	96.8 (60/62)
Yes	3.2 (2/62)
Findings suggestive of BV on Papanicolaou smear (% (n/N))	
No	64.5 (40 /62)
Yes	35.5 (22/62)

Table 1 (*continued*)

Characteristic	All participants ($N = 62$)
Cigarette use (% (n/N))	
Never smoked	79.0 (49/62)
Ex-smoker	1.6 (1/62)
Current smoker	19.4 (12/62)

Notes.

HPV, human papillomavirus; ASCUS, atypical cells of undetermined significance; LSIL, low-grade squamous intraepithelial lesion; HSIL, high-grade squamous intraepithelial lesion; BV, bacterial vaginosis.

Continuous variables are expressed as medians with interquartile ranges (IQRs, at 25th and 75th percentiles).

[a]Data was not available on the age at sexual debut for two women, lifetime number of sexual partners of two women and number of sexual acts with study partner in the last month of six women.

[b]Injectable progestin contraceptives.

[c]The identity of the oral pills (whether oestrogen or progestin or combination) was unknown.

(*Gardnerella*, *Prevotella*, *Sneathia*, *Shuttleworthia*, *Clostridium*, *Atopobium*, *Dialister*, and a consortium of low-abundant bacteria) and very low numbers of *Lactobacillus*, including *L. iners* and three unclassified *Lactobacillus* (*Lactobacillus.1*, *Lactobacillus.3*, and *Lactobacillus.4*). While there was a continuum of relative abundances of the bacterial taxa in CST III, four sub-clusters were evident (Fig. 1), three sub-clusters were dominated by *Shuttleworthia* ($n = 4$ women, 11.4%), *Gardnerella* ($n = 8$, 22.9%), and *Sneathia* ($n = 5$, 14.3%), and one with mixed taxa ($n = 18$, 51.4%). Next, the metadata for the women in each CST were compared.

Comparison of the community state types by participants' metadata

The demographic, sexual, smoking, and clinical characteristics of the women assigned to each of the three CSTs are shown in Table 2. The metadata for women in CST I and CST III were compared, while CST II was excluded from statistical comparisons due to the small sample size.

A significantly greater number of women with cervical microbiota from CST I reported hormonal contraceptive use compared to women with CST III (15/22 (68.2%) versus 9/31 (29%), $p = 0.005$). Findings suggestive of BV on smears were significantly more frequent ($p = 0.001$) in women with CST III than CST I. The other participants' variables, including HPV status, were not significantly different between the women in CST I and III.

Comparison of alpha diversity across CSTs, BV, HPV, and hormonal contraceptive use

Alpha diversity in the cervical microbiota was estimated using a variety of indices, including Simpson, Dominance, Shannon Diversity and Shannon Equitability (Fig. 2). A higher Dominance, Shannon, and Shannon Equitability index value, and lower Simpson index value designates greater alpha diversity. Based on rarefaction curves of alpha diversity metrics (Shannon and Simpson indices, Fig. S2), we chose 5,000 reads per sample as a sufficient subsampling depth to accurately assess microbial diversity.

When grouped by CST (Fig. 2A), CST I (*L. iners*-dominated group) and CST III (diverse group) were significantly different for all the alpha diversity indices ($p < 0.0001$). Thus, bacterial diversity in CST III was significantly greater than CST I.
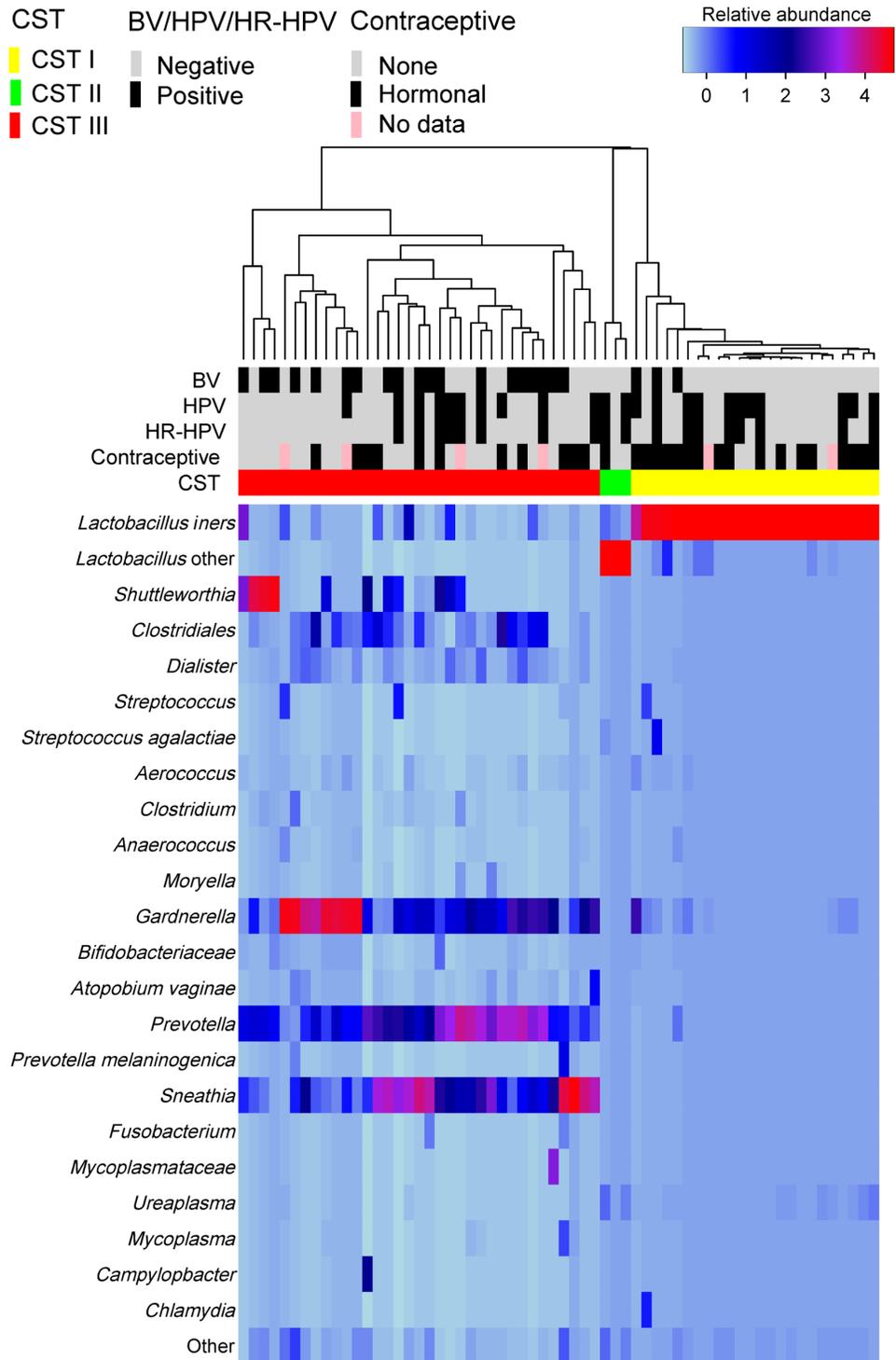
Figure 1 **Heatmap of the relative abundances of bacterial taxa in the cervical microbiota of 62 Black South African women.** The rows represent the bacterial taxa and columns the samples. The 23 most abundant taxa are displayed, with less abundant taxa grouped as "Other". (continued on next page...)

Figure 1 (…continued)
The names of the bacteria are presented at the deepest taxonomic level that they were assigned. The dendrogram depicts the average linkage hierarchical clustering of the cervical microbiota based on the Bray-Curtis dissimilarity. The cervical microbiota community state types (CSTs), human papillomavirus (HPV) and high-risk human papillomavirus (HR-HPV) infection status, bacterial vaginosis (BV) findings and contraceptive usage of the women are indicated.
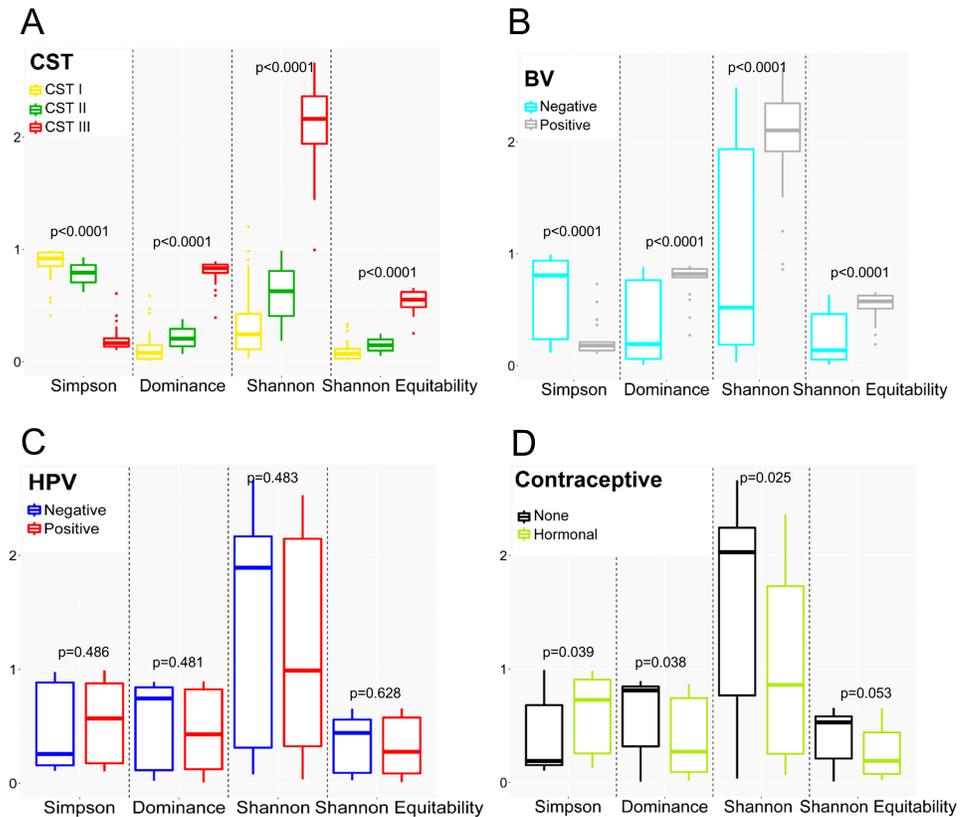
Figure 2 Alpha diversity measures of cervical microbiota. Comparison of the alpha diversity of the cervical microbiota grouped by: (A) Community state type (CST). (B) Bacterial vaginosis (BV) status. (C) Human papillomavirus (HPV) infection status. (D) Hormonal contraceptive usage. Women with missing information on hormonal contraception were excluded from this analysis. Each box plot is colour-coded according to the type of CST, HPV, BV, and hormonal contraception. In each plot, the box ranges from the first to the third quartile, with the median represented by the horizontal line. The whiskers extend to the smallest and the largest non-outliers and outliers represented by the dots.

The diversity of the cervical microbiota of women with findings suggestive of BV was significantly greater than that of women without findings of BV (Shannon index: 2.1 (1.8–2.4) versus 0.5 (0.2–1.9), $p < 0.0001$, Fig. 2B. Alpha diversity was also shown to be significant by Simpson, Dominance, Shannon, and Shannon Equitability metrics.

When grouped by HPV status (Fig. 2C), no significant difference in alpha diversity was observed between the HPV-negative and HPV-positive groups, with Shannon index of 1.9 (0.3–2.2) and 1.0 (0.3–2.2), $p = 0.483$, respectively. Alpha diversity was also shown to be significant by Simpson, Dominance, Shannon, and Shannon Equitability metrics.

Table 2 Clinical, demographic, sociobehavioural, and microbiological characteristics of the women by cervical community state type.

Characteristic	CST I (N = 24, 38.7%)	CST II (N = 3, 4.8%)	CST III (N = 35, 56.5%)	p-value[a]
Age (years)	30.5 (22.3–37.8)	31.0 (30.0–32.0)	35.0 (28.0–40.0)	0.203
HPV infection (% (n/N))				
Any HPV type	45.8 (11/24)	66.7 (2/3)	28.6 (10/35)	0.174
Any high-risk type	33.3 (8/24)	66.7 (2/3)	22.9 (8/35)	0.374
Single infection	25.0 (6/24)	66.7 (2/3)	20.0 (7/35)	0.301
Multiple infection	20.8 (5/24)	0.0 (0/3)	8.6 (3/35)	
HPV status at 6 month visit[e] (% (n/N))				
Negative	42.9 (6/14)	0.0 (0/2)	56.0 (14/25)	
Acquired	0.0 (0 /14)	0.0 (0/2)	16.0 (4/25)	0.073
Cleared	7.1 (1/14)	50.0 (1/2)	16.0 (4/25)	
Persisted	21.4 (3/14)	50.0 (1/2)	8.0 (2/25)	
Age at sexual debut (years)[d]	17.0 (16.0–18.0)	18.0 (18.0–19.0)	18.0 (17.0–19.0)	0.289
Lifetime number of sexual partners	2.0 (2.0–3.0)	2.0 (2.0–6.0)	2.0 (2.0–4.0)	0.711
Number of sex acts with study partner in last month[d]	2.0 (2.0–4.0)	2.0 (2.0–4.0)	2.0 (2.0–4.0)	0.445
Currently on hormonal contraceptives (% (n/N))	68.2 (15/22)	33.3 (1/3)	29.0 (9/31)	0.005
Depo-Provera[f]	50.0 (11/22)	33.3 (1/3)	19.4 (6/31)	0.019
Nonrethisterone enanthate[f]	18.2 (4/22)	0.0 (0/3)	3.2 (1/31)	0.147
Oral pills[g]	0.0 (0/22)	0.0 (0/3)	6.5 (2/31)	0.505
Cervical cytology (% (n/N))				
Normal	75.0 (18/24)	33.3 (1/3)	84.4 (27/32)	
ASCUS	4.2 (1/24)	33.3 (1/3)	6.3 (2/32)	0.536
LSIL	16.7 (4/24)	0.0 (0/3)	9.4 (3/32)	
HSIL	4.2 (1/24)	33.3 (1/3)	0.0 (0/32)	
Experienced vaginal discharge in last 6 months (% (n/N))	20.8 (5/24)	0.0 (0/3)	14.3 (5/35)	0.726
Experienced genital ulceration in last 6 months (% (n/N))	4.2 (1/24)	0.0 (0/3)	2.9 (1)	1.000
Positive for findings suggestive of BV on Papanicolaou smear (% (n/N))	12.5 (3/24)	0.0 (0/3)	54.3 (19/35)	0.001
Cigarette use (% (n/N))				
Never smoked	87.5 (21/24)	100.0 (3/3)	71.4 (25/35)	
Ex-smoker	4.2 (1/24)	0.0 (0/3)	0.0 (0/35)	0.091
Current smoker	8.3 (2/24)	0.0 (0/3)	28.6 (10/35)	
Prevalence (and mean relative abundance) of descriptive microbiological feature (bacteria)[b]				
L. iners	100.0 (91.7)	100.0 (3.2)	100.0 (2.9)	<0.0001
L. coleohominis	12.5 (<0.1)	0.0 (0.0)	0.0 (0.0)	0.062
L. mucosae	4.2 (<0.1)	0.0 (0.0)	0.0 (0.0)	0.407
L. ruminis	12.5 (<0.1)	0.0 (0.0)	0.0 (0.0)	0.062
Lactobacillus.1	100.0 (0.3)	100.0 (0.1)	14.3 (<0.1)	<0.000[c]
Lactobacillus.2	37.5 (0.1)	33.3 (0.5)	0.0 (0.0)	0.0001[c]
Lactobacillus.3	45.8 (1.1)	33.3 (<0.1)	5.7 (<0.1)	<0.0001[c]
Lactobacillus.4	37.5 (0.1)	100.0 (87.8)	45.7 (<0.1)	0.648

Characteristic	CST I (N = 24, 38.7%)	CST II (N = 3, 4.8%)	CST III (N = 35, 56.5%)	p-value[a]
Clostridiales	29.2 (<0.1)	33.3 (<0.1)	94.3 (6.0)	<0.0001[c]
Dialister	33.3 (<0.1)	33.3 (<0.1)	100.0 (2.3)	<0.0001[c]
Gardnerella	91.6 (2.5)	33.3 (<0.1)	100.0 (25.3)	<0.0001
Prevotella	75.0 (0.3)	100.0 (0.2)	100.0 (22.5)	<0.0001[c]
Shuttleworthia	33.3 (<0.1)	66.7 (<0.1)	54.3 (9.7)	<0.0001
Sneathia	75.0 (<0.1)	66.7 (<0.1)	97.1 (22.7)	<0.0001[c]
Alpha diversity				
Simpson index	0.9 (0.8–1.0)	0.8 (0.6–0.9)	0.2 (0.1–0.2)	<0.0001
Dominance index	0.1 (0.0–0.2)	0.2 (0.1–0.4)	0.8 (0.8–0.9)	<0.0001
Shannon index	0.2 (0.1–0.5)	0.6 (0.2–1.0)	2.2 (1.9–2.4)	<0.0001
Shannon Equitability index	0.1 (0.0–0.1)	0.1 (0.1–0.3)	0.6 (0.5–0.6)	<0.0001

Notes.

HPV, human papillomavirus; ASCUS, atypical cells of undetermined significance; LSIL, low-grade squamous intraepithelial lesion; HSIL, high-grade squamous intraepithelial lesion; BV, bacterial vaginosis; CST, community state type.

[a] p-values are shown for the comparison of CST I and CST III. Associations of continuous variables (expressed as medians with interquartile ranges (IQRs, at 25th and 75th percentiles)) and categorical variables were computed by Mann-Whitney unpaired and Chi-square/Fishers exact tests, respectively. CST II was excluded from the statistical analyses due to the low sample number (n = 3; 4.8%). Significant p-values (<0.05) are shown in bold.

[b] p-values are for differences in relative abundances.

[c] Prevalences were significantly different (*Lactobacillus.1* ($p < 0.0001$), *Lactobacillus.2* ($p = 0.0001$), *Lactobacillus.3* ($p = 0.0003$), *Clostridiales* ($p < 0.0001$), *Dialister* ($p < 0.0001$), *Prevotella* ($p = 0.003$), *Sneathia* ($p = 0.015$)).

[d] Data was not available on the age at sexual debut for two women (one CS T I and one CST III) and number of sexual acts wit h study partner in the last month for two women (one CST I and one CST III).

[e] Data was available for 41 women only. Five of these women (four CST I and one CST III) had complex HPV infection patterns, which were a combination of either cleared and acquired (three women) or cleared and persistent (two women) infections with specific HPV genotypes.

[f] Injectable progestin contraceptives.

[g] The identity of the oral pills (whether oestrogen or progestin or combination) was unknown.

Lastly, hormonal contraceptive users (Shannon index: 0.9 (0.2–1.9)) had a lower microbial diversity than non-users (2.3 (0.6–2.3)), $p = 0.025$ (Fig. 2D). When stratified by the type of contraceptive used, the diversities based on Shannon index were significantly different ($p = 0.028$); microbial diversity of Depo-Provera (0.7 (0.2–1.7)) and norethisterone enanthate (0.5 (0.2–1.4)) was significantly lower than that of non-users (2.0 (0.6–2.3)) (Fig. S3).

Comparison of beta diversity across CSTs, BV, HPV, and hormonal contraceptive use

Beta diversity analysis (PCoA of weighted UniFrac distances) of the 62 samples showed that each of the established CST (I-III) represented a highly distinct bacterial community, $p = 0.001$ (Fig. 3A). This result was supported by the Jackknife replicates that were used to estimate the uncertainty in PCoA plots and hierarchical clustering of the cervical microbiota. The majority of the samples with *Lactobacillus*-dominated cervical microbiota (CST I and CST II, 25/27, 92.6%) exclusively clustered together in the upper right quadrant (Fig. 3A). Two samples from CST I did not cluster with this group likely due to presence of other bacterial taxa, e.g., *Gardnerella*, *Prevotella*, and *Aerococcus*, in these cervical microbiota. These samples were from women with BV.

Further, beta diversity analysis showed that the clustering of the samples was dependent on the findings suggestive of BV, $p = 0.001$ (Fig. 3B). The 25 samples that clustered together
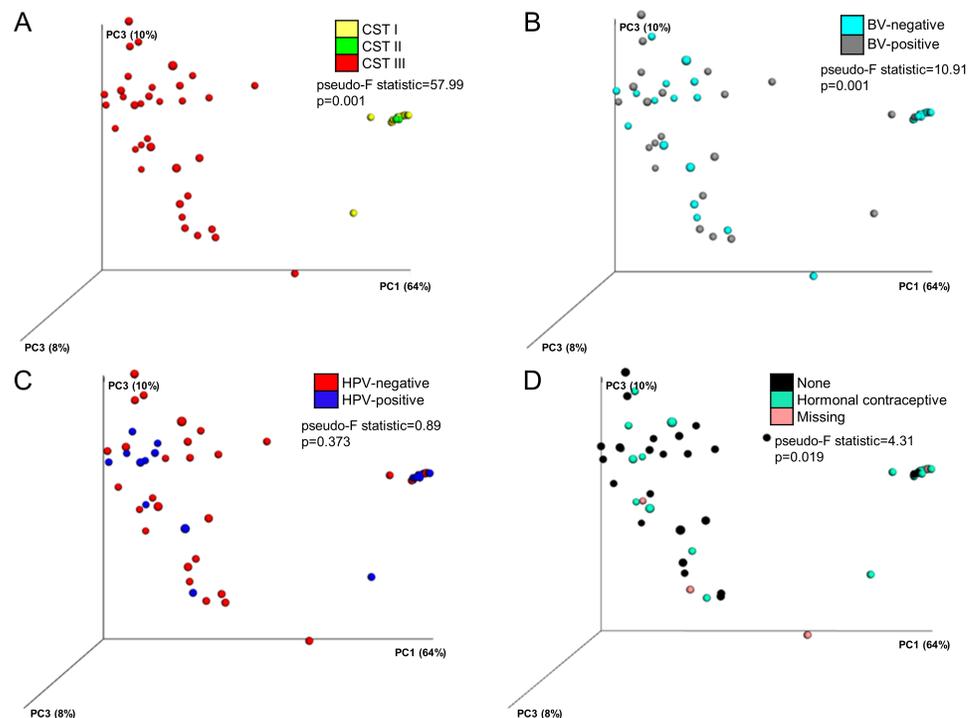
Figure 3 Beta diversity of the cervical microbiota. Principal Coordinates Analysis (PCoA) plots of the weighted UniFrac distances of the cervical microbiota coloured according to: (A) Community state type (CST). (B) Bacterial vaginosis (BV) status. (C) Human papillomavirus (HPV) infection status. (D) Hormonal contraception usage. The first three principal coordinate (PC) axes and the percentage variation explained by each (PC1: 64%, PC2: 10%, and PC3: 8%) are shown. Each solid point represents a bacterial community.

in the upper right quadrant consisted mostly of women without findings suggestive BV (22/25, 88.0%). Samples from women with findings suggestive of BV were spread over a greater area in the plot due to their high and varying bacterial diversity.

The weighted UniFrac distances of the cervical microbiota showed that there was no apparent influence of HPV infection on beta diversity, $p = 0.373$ (Fig. 3C). Of the samples that clustered together, 48.0% (12/25) were HPV-positive.

The majority of the samples from women on hormonal contraception (16/25, 64.0%), mostly Depo-Provera (12/18, 75.0%), clustered together in the upper right quadrant, $p = 0.019$ (Fig. 3D). Whereas 80.0% (4/5) of the samples from women on norethisterone enanthate clustered together with the 75.0% samples from women on Depo-Provera, all the samples from women on oral contraceptives (2/2, 100.0%) did not ($p = 0.018$).

Co-occurrence and co-exclusion patterns of cervical bacterial OTUs

Pairwise correlations were calculated for all pairs of OTUs identified in the cervical microbiota. A correlation matrix of the pairwise correlations between the 60 most abundant cervical bacterial OTUs is shown in Fig. 4.
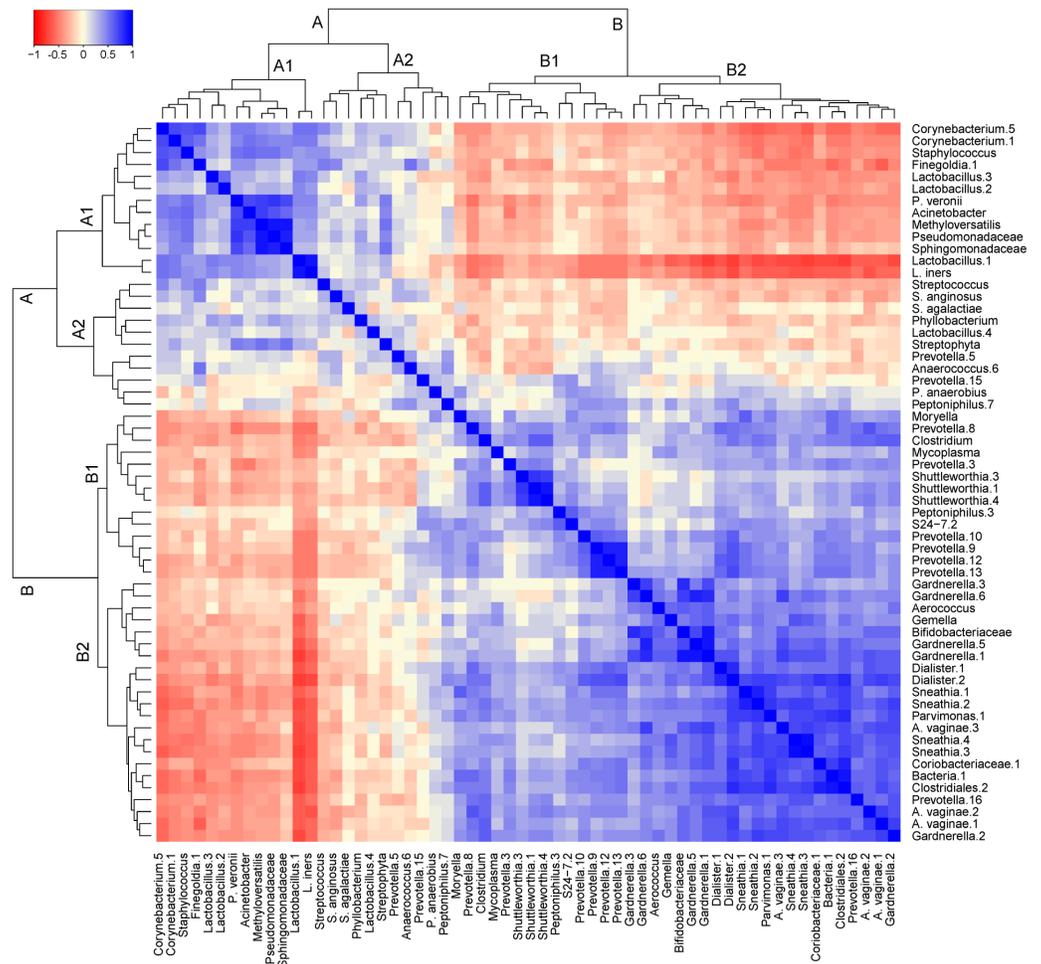
Figure 4 **Correlogram of 60 cervical bacterial OTUs showing co-occurrence and co-exclusion patterns.** Spearman's rank correlations between OTU counts were calculated in metagenomeSeq and the samples clustered. The correlation coefficients range from −1 (red; incompatibilities, co-exclusions, or oppositional interactions) to +1 (blue; symbiotic, mutualistic, or co-occurrence interactions.

    From the dendrogram, two major bacterial correlation clusters, Cluster-A (mostly with OTUs classified as *Lactobacillus* and *Streptococcus*) and Cluster-B (mostly with BV-associated bacteria) were observed. Each of these clusters had two sub-clusters: Cluster-A1 and Cluster-A2 for Cluster-A, and Cluster-B1 and Cluster-B2 for Cluster-B. OTUs in Cluster-A had an inverse correlation with OTUs in Cluster-B. For example, *Lactobacillus* OTUs (in Cluster-A) had negative correlations with *Gardnerella* and *Prevotella* OTUs (in Cluster-B). OTUs in the same Cluster-A1 had stronger positive correlations with one another than with OTUs in another sub-cluster from the same cluster (e.g., Cluster-A2). There was some overlap in the interaction of bacteria in Cluster-A2 and Cluster-B1. Interactions between these sub-clusters appeared very low to moderate. We noted that strong positive correlations were very common between phylogenetically related bacterial OTUs, e.g., *Lactobacillus* spp., but the extent of these interactions varied.

DISCUSSION

Using a culture-independent analysis of cervical microbiota of 62 reproductive HIV-seronegative Black women, we identified three CSTs (CST I: dominated by *L. iners*, CST II: dominated by an unclassified *Lactobacillus* OTU, and CST III: diverse and heterogeneous cervical microbiota) and found a positive association of hormonal contraception (mostly progestin-based) with CST I.

Among the women with *Lactobacillus* dominance in our cohort, CST I (*L. iners*-dominated) was the most prevalent (89%). *L. iners*-dominated CVM are the most prevalent CSTs with *Lactobacillus* dominance among women of African ancestry (*Borgdorff et al., 2017*; *Fettweis et al., 2014*; *Ravel et al., 2011*; *Zhou et al., 2007*), including those in sub-Saharan Africa (prevalence: 42–88%) (*Anahtar et al., 2015*; *Balle et al., 2018*; *Borgdorff et al., 2014*; *Dareng et al., 2016*; *Gautam et al., 2015*; *Lennard et al., 2018*; *Onywera et al., 2019*; *Wessels et al., 2017*). The variations in the prevalence of *L. iners* predominance (42–89%) could be attributed to the differences in anatomical sample type (whether cervical or vaginal) (*Balle et al., 2018*; *Huang et al., 2015*; *Kim et al., 2009*), sampling technique (*Kim et al., 2009*), behaviour (*Borgdorff et al., 2017*; *Brotman et al., 2014a*; *Vodstrcil et al., 2017*; *Wessels et al., 2017*), and genetics within the Black population (*Borgdorff et al., 2017*; *Tishkoff & Williams, 2002*). For instance, a study on a Kenyan cohort found that non-sex workers were more likely to have *L. iners*-dominated microbiota compared to female sex workers, FSWs (women who engaged in high-risk sexual behaviour) (*Wessels et al., 2017*). Overall, all the women (100.0%) in our study had detectable *L. iners*. This finding is qualitatively congruent with those of culture-independent studies (*Borgdorff et al., 2014*; *Ravel et al., 2011*) but contrasts those reported in cultivation studies on premenopausal South African women (36–75%) (*Damelin et al., 2011*; *Pendharkar et al., 2013*). This dissimilarity could be due to difficulty in culturing *L. iners* (*Lamont et al., 2011*; *Srinivasan et al., 2016*), the higher sensitivity of deep sequencing technology (*Fettweis et al., 2012*; *Nielsen et al., 2007*), and/or ability of DNA-based identification methods (e.g., PCR) to detect relic DNA (from non-viable or dead bacteria) besides DNA from viable bacteria (*Nielsen et al., 2007*). It is believed that the detectability of *Lactobacillus* spp. in all the cervical microbiota attests that the production of lactic acid is conceivably conserved in all microbiota (*Ravel et al., 2011*), thus pinpointing it as the core acidifier (*Boskey et al., 1999*).

The role of *L. iners* in the cervicovaginal health is however unclear (*Petrova et al., 2017*). Unlike many *Lactobacillus* spp., *L. iners* can occur with BV-associated bacteria (*Borgdorff et al., 2017*; *Gautam et al., 2015*; *Petrova et al., 2017*; *Srinivasan et al., 2012*) and can at times enhance their adhesion to cervical epithelium (*Castro et al., 2013*). Additionally, it has been consistently isolated from women with and without vaginal syndromes, intermediate flora (*Damelin et al., 2011*; *Pendharkar et al., 2013*; *Petrova et al., 2017*; *Srinivasan et al., 2016*; *Zozaya-Hinchliffe et al., 2010*), or women with CSTs transitioning to healthy or dysbiotic states (*Gajer et al., 2012*; *Petrova et al., 2017*). Growth of *L. iners* in BV-associated environment could be due to its inefficient colonization resistance to opportunistic and pathogenic bacteria (*Pendharkar et al., 2013*; *Zhou et al., 2007*) or better tolerant and survival phenotypes even in perturbed milieus (*Pendharkar et al., 2013*; *Zozaya-Hinchliffe*

et al., 2010). There is compelling omics evidence supporting the second explanation. Genomics have suggested that *L. iners* underwent rapid evolutionary events that endowed it with competitive and specialized adaptation capabilities even in dysbiotic milieu (*Macklaim et al., 2010*). Meta-transcriptomics have strengthened these facts, demonstrating that *L. iners* is able to differentially express over 10% of its genome in order to survive in dysbiotic state (*Macklaim et al., 2013*). *L .iners* can predispose women to an aberrant microbiota or BV (*Verstraelen et al., 2009*) and has been associated with STIs (*Borgdorff et al., 2014*; *Brotman et al., 2014b*; *Van Houdt et al., 2018*). In contrast to this detrimental outcome, *L. iners* can interfere with *G. vaginalis* biofilm assembly (*Saunders et al., 2007*), thereby restoring a healthy CVM. More exploratory studies are therefore needed to characterize cervicovaginal *L. iners* since it is currently believed that it has clonal variants that may have different roles in health, dysbiosis, and disease (*Petrova et al., 2017*).

In our study, about 5% of the women had CVM dominated by an unclassified *Lactobacillus* sp. (CST II). Similar to a previous study that used a similar methodology (*Roesch et al., 2017*), the V4 hypervariable region of the 16S rRNA gene did not allow us to achieve a deeper taxonomic discrimination of the *Lactobacillus* in CST II. Therefore, we could not ascertain whether the low prevalent CST II was one of the commonly established CST with *Lactobacillus* (*L. crispatus*, *L. gasseri*, or *L. jensenii*) dominance as found elsewhere (*Ravel et al., 2011*). Generally, *L. crispatus*, *L. gasseri*, and *L. jensenii* are often less common and less abundant in Black women compared to White women (*Balle et al., 2018*; *Borgdorff et al., 2014*; *Fettweis et al., 2014*; *Lennard et al., 2018*; *Ravel et al., 2011*; *Zhou et al., 2007*), with *L. crispatus* recently found to be less abundant in cervical samples (similar to our study's) compared to lateral vaginal wall samples (*Balle et al., 2018*). While we detected other *Lactobacillus* spp. such as *L. coleohominis*, *L. mucosae*, and *L. ruminis* that have been uncovered from premenopausal South African women (*Damelin et al., 2011*; *Pendharkar et al., 2013*), we also confirmed that microbiota with approximately equal dominance of two or more *Lactobacillus* spp. are absent or underrepresented in Black women (*Zhou et al., 2007*).

The prevalence of diverse and heterogeneous group, CST III (57%), was higher than has been documented in non-Black women (*Borgdorff et al., 2017*; *Fettweis et al., 2014*; *Ravel et al., 2011*; *Zhou et al., 2007*) and intermediate to what has been recently reported among Black South African women (47–64%) (*Anahtar et al., 2015*; *Lennard et al., 2018*; *Onywera et al., 2019*). CST III exhibited four intracluster variations: CST III-*Shuttleworthia*, CST III-*Gardnerella*, CST III-*Sneathia*, and CST III-mixed, which lacked a clear dominance. The observation of *Shuttleworthia* should be treated with scepticism, it is perhaps a misclassification of BV-associated bacterium-1 (BVAB-1) (*Oakley et al., 2008*). *Gardnerella*-dominated microbiota have been identified in African American (*Zhou et al., 2007*), Black South African (*Anahtar et al., 2015*), African Surinamese, and Ghanaian (*Borgdorff et al., 2017*) women. CST III was associated with BV, thus, confirming earlier findings (*Onywera et al., 2019*; *Ravel et al., 2011*; *Srinivasan et al., 2012*; *Zozaya-Hinchliffe et al., 2010*). It is important to point out that diversity of CVM as observed by 16S rRNA sequencing may not always be associated with BV as diagnosed by laboratory tests. This was corroborated in a study by Wessels and colleagues (*2017*) that found that a majority of FSW (74%) without

BV (Nugent score: 0–3) had highly diverse vaginal communities (as captured by 16S rRNA sequencing) (*Wessels et al., 2017*).

While this study was not specifically designed to assess the impact of hormonal contraceptives on the cervical microbiota, we did observe that hormonal contraceptive use (mostly progestin: at least 92%) was correlated with *L. iners* dominance of the cervical microbiota in reproductive-age Black South African women. At the moment, there is a wide heterogeneity of results regarding the effect of hormonal contraception on the composition and diversity of the CVM. Such results include (i) increased detectability of *Lactobacillus fermentum* among Pakistani woman on oral contraception (*Kazi, Saleem & Kazi, 2012*), (ii) higher abundances of *Lactobacillus* spp. (*L. crispatus* and *L. jensenii*) and *L. iners* among US women on oral and progestin-based contraception, respectively (*Brooks et al., 2017*), (iii) no effect on *L. iners* quantities, reduced total bacterial load and *G. vaginalis* following progestin-based initiation among Kenyan women (*Roxby et al., 2016*), (iv) decreased levels of *L. iners* but not of the other beneficial *Lactobacillus* spp. (*L. crispatus*, *L. gasseri*, and *L. jensenii*) in response to progestin-based contraception among Zimbabwean women (*Achilles et al., 2018*), (v) reduced lactobacilli following long-term use of progestin-based contraception among US women (*Mitchell et al., 2014*), and so on. We further noted that women using hormonal contraceptives (mostly progestin) had a lower bacterial (alpha) diversity relative to non-users. Previous studies have found either no discernible (*Birse et al., 2017*) or positive associations between progestin-based contraception and the level of bacterial diversity (*Brooks et al., 2017*; *Jespers et al., 2017*; *Yang et al., 2019*). The differential effects of hormonal contraception, particularly progestin-based, could be governed by host genetics (*Yang et al., 2019*) and differences in study methodologies. Although current literature suggests that hormonal contraceptives (oestrogen- and/or progestin-based) reduces the risk of BV-associated bacteria (*Brooks et al., 2017*), BV-like microbiota (*Borgdorff et al., 2017*) and BV (*Abbai, Reddy & Ramjee, 2016*; *Vodstrcil et al., 2013*), with progestin-based contraceptives reducing BV risk by 18–30% (*Van de Wijgert et al., 2013*), more studies are still needed to fully understand the microbiological consequences of progestin-induced amenorrhea and systemic hypoestrogenism in the context of reproductive health (*Miller et al., 2000*).

In the present study, except for BV and hormonal contraception, we did not find any association of cervical microbiota with the other participant variables. For example, we observed no significant difference in prevalent HPV between *L. iners* dominated and diverse microbiota, a finding that mirrors previous studies (*Borgdorff et al., 2014*; *Brotman et al., 2014b*; *Onywera et al., 2019*). There may be no association or, more likely, this was due to methodological shortcomings in our study, including the relatively small sample size.

Finally, our results on correlational analyses provided evidence for synergism (co-occurrence) and antagonism (co-exclusion) of the cervicovaginal bacterial communities. These ecological patterns have been elucidated to act concurrently to structure the microbiota compositions (*Zelezniak et al., 2015*). Positive correlations suggest niche sharing, similar nutrient requirements and or metabolic co-dependencies. Co-occurrence patterns, for example, of *Gardnerella* with *Prevotella* and other BV-associated bacteria parallels previous findings (*Ravel et al., 2011*). Moreover, phylogenetically related bacterial

OTUs such as *Prevotella* spp. had strong positive correlations, which may be due to the high-level of resource overlap (*Zelezniak et al., 2015*). Variations in positive correlations of more closely phylogenetically-related bacteria like the different *Atopobium* spp., or *Gardnerella* spp., illuminates the existence of diverse bacterial genetic profiles at species level (*De Backer et al., 2006*; *Eren et al., 2011*), plausibly with different virulent competencies. Strains of *Atopobium* spp. and *Gardnerella* spp. have been demonstrated to have dissimilar phenotypic behaviours in cervicovaginal health and disease (*De Backer et al., 2006*; *Swidsinski et al., 2010*). Inverse correlations such as those observed between *Lactobacillus* and BV-associated bacterial OTUs in our study and others (*Anahtar et al., 2015*; *Castro et al., 2013*; *Ravel et al., 2011*; *Srinivasan et al., 2012*; *Srinivasan et al., 2010*), are indications for niche filtering and/or competition for growth nutrients (*Srinivasan et al., 2012*; *Zelezniak et al., 2015*). As stated by Ravel and co-workers (*2011*), the precise relevance of these bacterial positive and negative interactions remains undefined; therefore, subject to further investigations.

Even though the present study broadens our knowledge about the ethnic differences in the composition of cervical microbiota and associations of particular microbiota with clinical and behavioural characteristics, a few limitations arising from this retrospective cross-sectional study should be noted. First, while we found an association between hormonal contraception (mostly progestin) and *L. iners* dominance, we did not adjust our analysis for potential confounders such as age, time of sampling (with regards to hormonal/menstrual cycle stage and length of time since last dose of hormonal contraception), length of time on hormonal contraception, sexual behaviour, STIs, grouping all forms of hormonal contraceptives together (as they might be exerting different effects on cervical microbiota), and vaginal disorders (e.g., BV and aerobic vaginitis) to name a few. Secondly, small sample size in some groups limited additional comparisons. In addition, we might have inadequately diagnosed BV since we used cervical samples and a non-standard approach (Pap smear) to diagnose BV. It is believed that BV is best diagnosed with vaginal instead of cervical samples (*Hillier, 1993*). Pap smear has been demonstrated not to perform well (specificity: 93–94% and sensitivity: 43–49%) when compared to Gram stain diagnosis (*Greene, Kuehl & Allen, 2000*; *Tokyol et al., 2004*). Some investigators consider techniques that primarily rely on the presence of *Gardnerella* to diagnose BV to be inappropriate. (*Dols et al., 2011*). However, the high specificity of Pap smear suggests that "it may be an adequate diagnostic criterion when it is positive" (*Tokyol et al., 2004*). Lastly, failure to confidently assign species names to some bacteria impeded accurate comparison of our results to other published studies. Future studies would be of more benefit if all these limitations are addressed.

CONCLUSIONS

A majority of the reproductive-age HIV-seronegative Black South African women (57%) had cervical microbiota not dominated by *Lactobacillus*, the bacteria assumed to constitute a healthy cervical microbiota. These cervical microbiota were associated with findings suggestive of BV. It has been speculated that such cervical microbiota may be a contributing factor to the high burden of HIV and HPV infections among Black women (*Zhou et al.,*

2007). Not all women (46%) with non-*Lactobacillus*-dominated cervical microbiota had findings suggestive of BV. Hence, additional studies are needed to examine whether these cervical microbiota signify abnormal, intermediate or variant states of health in Black women. The association of hormonal contraceptive (mostly progestin) use with *L. iners* dominance merits further investigation as there is still paucity of studies, uncertainty and controversy surrounding this topic.

ACKNOWLEDGEMENTS

We are grateful to all the study participants and personnel involved in the HPV Couples Cohort Study; with special mention to Prof. Jennifer Moodley, one of the co-principal investigators. Our computational analyses were performed using facilities provided by the University of Cape Town's ICTS High Performance Computing team: http://hpc.uct.ac.za/. We thank two members of the Computational Biology (CBIO) Division, University of Cape Town—South Africa: Gerrit Botha for the metagenomics training and advice, and Brejnev Muhire, for the Python language advice. We also appreciate Sarah Bonnin of the Bioinformatics Core Facility at the Center for Genomic Regulation (CRG)—Spain, for her immense guidance in R programming.

Funding

This work is based on the research supported by the National Research Foundation (NRF) of South Africa (Grant Number: 64815), the Poliomyelitis Research Foundation (PRF), the Cancer Association of South Africa (CANSA), the University of Cape Town (UCT) Research Incentive Scheme, and the UCT Cancer Research Initiative. Harris Onywera was a recipient of the UCT International and Refugee Students' Scholarship and the UCT postgraduate publication incentive (PPI). There was no additional external funding received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
National Research Foundation (NRF) of South Africa: 64815.
Poliomyelitis Research Foundation (PRF).
the Cancer Association of South Africa (CANSA).
the University of Cape Town (UCT) Research Incentive Scheme, and the UCT Cancer Research Initiative.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Harris Onywera conceived and designed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft, optimized the bioinformatics pipeline.
- Anna-Lise Williamson conceived and designed the experiments, contributed reagents/materials/analysis tools, authored or reviewed drafts of the paper, approved the final draft.
- Zizipho Z.A. Mbulawa and David Coetzee contributed reagents/materials/analysis tools, authored or reviewed drafts of the paper, approved the final draft.
- Tracy L. Meiring conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.

Human Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

This study was approved by the Human Research Ethics Committee of the University of Cape Town, South Africa (references 258/2006 and 580/2014).

Data Availability

Data is available at NCBI SRA, accession numbers: SRX4103412, SRX4103413, SRX4103414, SRX4103415, SRX4103416, PRJNA472137.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.7488#supplemental-information.

REFERENCES

Abbai NS, Reddy T, Ramjee G. 2016. Prevalent bacterial vaginosis infection— a risk factor for incident sexually transmitted infections in women in Durban, South Africa. *International Journal of STD and AIDS* **27**:1283–1288 DOI 10.1177/0956462415616038.

Achilles SL, Austin MN, Meyn LA, Mhlanga F, Chirenje ZM, Hillier SL. 2018. Impact of contraceptive initiation on vaginal microbiota. *American Journal of Obstetrics and Gynecology* **218(6)**:622.e1–622.e10 DOI 10.1016/j.ajog.2018.02.017.

Anahtar MN, Byrne EH, Doherty KE, Bowman BA, Yamamoto HS, Soumillon M, Padavattan N, Ismail N, Moodley A, Sabatini ME, Ghebremichael MS, Nusbaum C, Huttenhower C, Virgin HW, Ndung'u T, Dong KL, Walker BD, Fichorova RN, Kwon DS. 2015. Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. *Immunity* **42**:965–976 DOI 10.1016/j.immuni.2015.04.019.

Anderson MJ. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* **26**:32–46 DOI 10.1111/j.1442-9993.2001.01070.pp.x.

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. *Available at http://www.bioinformatics.babraham.ac.uk/projects/fastqc/* (accessed on 5 April 2018).

Balle C, Lennard K, Dabee S, Barnabas SL, Jaumdally SZ, Gasper MA, Maseko V, Mbulawa ZZA, Williamson AL, Bekker LG, Lewis DA, Passmore JS, Jaspan HB. 2018. Endocervical and vaginal microbiota in South African adolescents with asymptomatic *Chlamydia trachomatis* infection. *Scientific Reports* **8**:11109 DOI 10.1038/s41598-018-29320-x.

Beamer MA, Austin MN, Avolia HA, Meyn LA, Bunge KE, Hillier SL. 2017. Bacterial species colonizing the vagina of healthy women are not associated with race. *Anaerobe* **45**:40–43 DOI 10.1016/j.anaerobe.2017.02.020.

Birse KD, Romas LM, Guthrie BL, Nilsson P, Bosire R, Kiarie J, Farquhar C, Broliden K, Burgener AD. 2017. Genital injury signatures and microbiome alterations associated with depot medroxyprogesterone acetate usage and intravaginal drying practices. *Journal of Infectious Diseases* **215(4)**:590–598 DOI 10.1093/infdis/jiw590.

Borgdorff H, Tsivtsivadze E, Verhelst R, Marzorati M, Jurriaans S, Ndayisaba GF, Schuren FH, Van de Wijgert JH. 2014. *Lactobacillus*-dominated cervicovaginal microbiota associated with reduced HIV/STI prevalence and genital HIV viral load in African women. *ISME Journal* **8**:1–3 DOI 10.1038/ismej.2013.176.

Borgdorff H, Van der Veer C, Van Houdt R, Alberts CJ, De Vries HJ, Bruisten SM, Snijder MB, Prins M, Geerlings SE, Schim van der Loeff MF, Van de Wijgert J. 2017. The association between ethnicity and vaginal microbiota composition in Amsterdam, the Netherlands. *PLOS ONE* **12(7)**:e0181135 DOI 10.1371/journal.pone.0181135.

Bosch FX, Lorincz A, Muñoz N, Meijer CJLM, Shah KV. 2002. The causal relation between human papillomavirus and cervical cancer. *Journal of Clinical Pathology* **55**:244–265 DOI 10.1136/jcp.55.4.244.

Boskey ER, Telsch KM, Whaley KJ, Moench TR, Cone RA. 1999. Acid production by vaginal flora *in vitro* is consistent with the rate and extent of vaginal acidification. *Infection and Immunity* **67**:5170–5175.

Brooks JP, Edwards DJ, Blithe DL, Fettweis JM, Serrano MG, Sheth NU, Strauss 3rd JF, Buck GA, Jefferson KK. 2017. Effects of combined oral contraceptives, depot medroxyprogesterone acetate and the levonorgestrel-releasing intrauterine system on the vaginal microbiome. *Contraception* **95**:405–413 DOI 10.1016/j.contraception.2016.11.006.

Brotman RM, He X, Gajer P, Fadrosh D, Sharma E, Mongodin EF, Ravel J, Glover ED, Rath JM. 2014a. Association between cigarette smoking and the vaginal microbiota: a pilot study. *BMC Infectious Diseases* **14**:471 DOI 10.1186/1471-2334-14-471.

Brotman RM, Shardell MD, Gajer P, Tracy JK, Zenilman JM, Ravel J, Gravitt PE. 2014b. Interplay between the temporal dynamics of the vaginal microbiota and

human papillomavirus detection. *Journal of Infectious Diseases* **210**:1723–1733 DOI 10.1093/infdis/jiu330.

Bruni L, Albero G, Serrano B, Mena M, Gómez D, Muñoz J, Bosch FX, DeSanjosé S. 2019. Human papillomavirus and related diseases in the world. Summary report 22 2019. ICO Information Centre on HPV and Cancer (HPV Information Centre).

Caporaso JG, Bittinger K, Bushman FD, De Santis TZ, Andersen GL, Knight R. 2010a. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* **26**:266–267 DOI 10.1093/bioinformatics/btp636.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich K, Gordon JI, Huttley GA, Kelley ST, Knights D, Jeremy E, Ley RE, Lozupone CA, McDonald D, Muegge BD, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsuneko T, Zaneveld J, Knight R. 2010b. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**:335–336 DOI 10.1038/nmeth.f.303.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences of the United States of America* **15**:4516–4522 DOI 10.1073/pnas.1000080107.

Castro J, Henriques A, Machado A, Henriques M, Jefferson KK, Cerca N. 2013. Reciprocal interference between *Lactobacillus* spp. and *Gardnerella vaginalis* on initial adherence to epithelial cells. *International Journal of Medical Sciences* **10**:1193–1198 DOI 10.7150/ijms.6304.

Damelin LH, Paximadis M, Mavri-Damelin D, Birkhead M, Lewis DA, Tiemessen CT. 2011. Identification of predominant culturable vaginal *Lactobacillus* species and associated bacteriophages from women with and without vaginal discharge syndrome in South Africa. *Journal of Medical Microbiology* **60**:180–183 DOI 10.1099/jmm.0.024463-0.

Dareng EO, Ma B, Famooto AO, Akarolo-Anthony SN, Offiong RA, Olaniyan O, Dakum PS, Wheeler CM, Fadrosh D, Yang H, Gajer P, Brotman RM, Ravel J, Adebamowo CA. 2016. Prevalent high-risk HPV infection and vaginal microbiota in Nigerian women. *Epidemiology and Infection* **144**:123–137 DOI 10.1017/S0950268815000965.

De Backer E, Verhelst R, Verstraelen H, Claeys G, Verschraegen G, Temmerman M, Vaneechoutte M. 2006. Antibiotic susceptibility of *Atopobium vaginae*. *BMC Infectious Diseases* **6**:51 DOI 10.1186/1471-2334-6-51.

De Santis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology* **72**:5069–5072 DOI 10.1128/AEM.03006-05.

Dols JAM, Smit PW, Kort R, Reid G, Schuren FHJ, Tempelman H, Bontekoe TR, Korporaal H, Boon ME. 2011. Microarray-based identification of clinically relevant vaginal bacteria in relation to bacterial vaginosis. *American Journal of Obstetrics and Gynecology* **204(4)**:305.e1-7 DOI 10.1016/j.ajog.2010.11.012.

Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* **10**:996–998 DOI 10.1038/nmeth.2604.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**:2194–2200 DOI 10.1093/bioinformatics/btr381.

Eren AM, Zozaya M, Taylor CM, Dowd SE, Martin DH, Ferris MJ. 2011. Exploring the diversity of *Gardnerella vaginalis* in the genitourinary tract microbiota of monogamous couples through subtle nucleotide variation. *PLOS ONE* **6(10)**:e26732 DOI 10.1371/journal.pone.0026732.

Fettweis JM, Brooks JP, Serrano MG, Sheth NU, Girerd PH, Edwards DJ, Strauss JF, Jefferson KK, Buck GA. 2014. Differences in vaginal microbiome in African American women versus women of European ancestry. *Microbiology* **160**:2272–2282 DOI 10.1099/mic.0.081034-0.

Fettweis JM, Serrano MG, Girerd PH, Jefferson K, Buck GA. 2012. A new era of the vaginal microbiome: advances using next-generation sequencing. *Chemistry & Biodiversity* **9**:965–976 DOI 10.1002/cbdv.201100359.

Gajer P, Brotman RM, Guoyun B, Sakamoto J, Schütte UME, Zhong X, Koenig SSK, Fu F, Ma Z, Zhou X, Abdo Z, Forney L, Ravel J. 2012. Temporal dynamics of the human vaginal microbiota. *Science Translational Medicine* **4**:132ra152–132ra152 DOI 10.1126/scitranslmed.3003605.

Gautam R, Borgdorff H, Jespers V, Francis SC, Verhelst R, Mwaura M, Delany-Moretlwe S, Ndayisaba G, Kyongo JK, Hardy L, Menten J, Crucitti T, Tsivtsi-vadze E, Schuren F, Van de Wijgert JH. 2015. Correlates of the molecular vaginal microbiota composition of African women. *BMC Infectious Diseases* **15**:86–86 DOI 10.1186/s12879-015-0831-1.

Greene JF, Kuehl TJ, Allen SR. 2000. The Papanicolaou smear: inadequate screening test for bacterial vaginosis during pregnancy. *American Journal of Obstetrics and Gynecology* **182**:1048–1049 DOI 10.1067/mob.2000.105394.

Hillier SL. 1993. Diagnostic microbiology of bacterial vaginosis. *American Journal of Obstetrics and Gynecology* **169**:455–459 DOI 10.1016/0002-9378(93)90340-O.

Huang YE, Wang Y, He Y, Ji Y, Wang LP, Sheng HF, Zhang M, Huang QT, Zhang DJ, Wu JJ, Zhong M, Zhou HW. 2015. Homogeneity of the vaginal microbiome at the cervix, posterior fornix, and vaginal canal in pregnant Chinese women. *Microbial Ecology* **69**:407–414 DOI 10.1007/s00248-014-0487-1.

Jespers V, Kyongo J, Joseph S, Hardy L, Cools P, Crucitti T, Mwaura M, Ndayisaba G, Delany-Moretlwe S, Buyze J, Vanham G, Van de Wijgert J. 2017. A longitudinal analysis of the vaginal microbiota and vaginal immune mediators in women from sub-Saharan Africa. *Scientific Reports* **7**:11974 DOI 10.1038/s41598-017-12198-6.

Kazi YF, Saleem S, Kazi N. 2012. Investigation of vaginal microbiota in sexually active women using hormonal contraceptives in Pakistan. *BMC Urology* **12**:22 DOI 10.1186/1471-2490-12-22.

Kenyon C, Colebunders R, Crucitti T. 2013. The global epidemiology of bacterial vaginosis: a systematic review. *American Journal of Obstetrics and Gynecology* **209**:505–523 DOI 10.1016/j.ajog.2013.05.006.

Kim TK, Thomas SM, Ho M, Sharma S, Reich CI, Frank JA, Yeater KM, Biggs DR, Nakamura N, Stumpf R, Leigh SR, Tapping RI, Blanke SR, Slauch JM, Gaskins HR, Weisbaum JS, Olsen GJ, Hoyer LL, Wilson BA. 2009. Heterogeneity of vaginal microbial communities within individuals. *Journal of Clinical Microbiology* **47**:1181–1189 DOI 10.1128/JCM.00854-08.

Koshiol J, Lindsay L, Pimenta JM, Poole C, Jenkins D, Smith JS. 2008. Persistent human papillomavirus infection and cervical neoplasia: a systematic review and meta-analysis. *American Journal of Epidemiology* **168**:123–137 DOI 10.1093/aje/kwn036.

Kurman RJ, Solomon D. 1994. *The Bethesda System for reporting cervical/vaginal cytologic diagnoses: definitions, criteria, and explanatory notes for terminology and specimen adequacy*. New York: Springer-Verlag New York Inc.

Lamont RF, Sobel JD, Akins RA, Hassan SS, Chaiworapongsa T, Pedro J, Romero R. 2011. The vaginal microbiome: new information about genital tract using molecular based techniques. *BJOG: An International Journal of Obstetrics & Gynaecology* **118**:533–549 DOI 10.1111/j.1471-0528.2010.02840.x.

Lennard K, Dabee S, Barnabas SL, Havyarimana E, Blakney A, Jaumdally SZ, Botha G, Mkhize NN, Bekker LG, Lewis DA, Gray G, Mulder N, Passmore JS, Jaspan HB. 2018. Microbial composition predicts genital tract inflammation and persistent bacterial vaginosis in adolescent South African women. *Infection and Immunity* **86(1)**:e00410-17 DOI 10.1128/IAI.00410-17.

Macklaim JM, Fernandes AD, Bella JMD, Hammond J-A, Reid G, Gloor GB. 2013. Comparative meta-RNA-seq of the vaginal microbiota and differential expression by *Lactobacillus iners* in health and dysbiosis. *Microbiome* **1**:12 DOI 10.1186/2049-2618-1-12.

Macklaim JM, Gloor GB, Anukam KC, Cribby S, Reid G. 2010. At the crossroads of vaginal health and disease, the genome sequence of *Lactobacillus iners* AB-1. *Proceedings of the National Academy of Sciences of the United States of America* **108**:205–209 DOI 10.1073/pnas.1000086107.

Mbulawa ZZA, Coetzee D, Marais DJ, Kamupira M, Zwane E, Allan B, Constant D, Moodley JR, Hoffman M, Williamson A-L. 2009. Genital human papillomavirus prevalence and human papillomavirus concordance in heterosexual couples are positively associated with human immunodeficiency virus coinfection. *Journal of Infectious Diseases* **199**:1514–1524 DOI 10.1086/598220.

Miller L, Patton DL, Meier A, Thwin SS, Hooton TM, Eschenbach DA. 2000. Depomedroxyprogesterone-induced hypoestrogenism and changes in vaginal flora and epithelium. *Obstetrics and Gynecology* **96**:431–439 DOI 10.1016/S0029-7844(00)00906-6.

Mitchell CM, McLemore L, Westerberg K, Astronomo R, Smythe K, Gardella C, Mack M, Magaret A, Patton D, Agnew K, McElrath MJ, Hladik F, Eschenbach D. 2014. Long-term effect of depot medroxyprogesterone acetate on vaginal microbiota,

epithelial thickness and HIV target cells. *Journal of Infectious Diseases* **210**:651–655 DOI 10.1093/infdis/jiu176.

Nardis C, Mosca L, Mastromarino P. 2013. Vaginal microbiota and viral sexually transmitted diseases. *Annali di Igiene* **25**:443–456 DOI 10.7416/ai.2013.1946.

Nielsen KM, Johnsen PJ, Bensasson D, Daffonchio D. 2007. Release and persistence of extracellular DNA in the environment. *Environmental Biosafety Research* **6**:37–53 DOI 10.1051/ebr:2007031.

Oakley BB, Fiedler TL, Marrazzo JM, Fredricks DN. 2008. Diversity of human vaginal bacterial communities and associations with clinically defined bacterial vaginosis. *Applied and Environmental Microbiology* **74**:4898–4909 DOI 10.1128/AEM.02884-07.

Onywera H, Williamson AL, Mbulawa ZZA, Coetzee D, Meiring TL. 2019. The cervical microbiota in reproductive-age South African women with and without human papillomavirus infection. *Papillomavirus Research* **7**:154–163 DOI 10.1016/j.pvr.2019.04.006.

Paulson JN, Stine OC, Bravo HC, Pop M. 2014. Robust methods for differential abundance analysis in marker gene surveys. *Nature Methods* **10**:1200–1202 DOI 10.1038/nmeth.2658.Robust.

Paulson JN, Talukder H, Pop M, Bravo HC. 2016. metagenomeSeq: statistical analysis for sparse high-throughput sequencing. Bioconductor package: 1.14.2. *Available at http://cbcb.umd.edu/software/metagenomeSeq*.

Pendharkar S, Magopane T, Larsson P-G, De Bruyn G, Gray GE, Hammarström L, Marcotte H. 2013. Identification and characterisation of vaginal lactobacilli from South African women. *BMC Infectious Diseases* **13**:43 DOI 10.1186/1471-2334-13-43.

Petrova MI, Lievens E, Malik S, Imholz N, Lebeer S. 2015. *Lactobacillus* species as biomarkers and agents that can promote various aspects of vaginal health. *Frontiers in Physiology* **6**:81 DOI 10.3389/fphys.2015.00081.

Petrova MI, Reid G, Vaneechoutte M, Lebeer S. 2017. *Lactobacillus iners*: friend or foe? *Trends in Microbiology* **25**:182–191 DOI 10.1016/j.tim.2016.11.007.

Philip D. 2003. Computer program review: VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* **14**:927–930 DOI 10.1111/j.1654-1103.2003.tb02228.x.

Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, Ault K, Peralta L, Forney LJ. 2011. Vaginal microbiome of reproductive-age women. *Proceedings of the National Academy of Sciences of the United States of America* **108**:4680–4687 DOI 10.1073/pnas.1002611107.

Roesch LF, Silveira RC, Corso AL, Dobbler PT, Mai V, Rojas BS, Laureano AM, Procianoy RS. 2017. Diversity and composition of vaginal microbiota of pregnant women at risk for transmitting Group B Streptococcus treated with intrapartum penicillin. *PLOS ONE* **12(2)**:e0169916 DOI 10.1371/journal.pone.0169916.

Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Nikita L, Galuppi M, Lamont RF, Chaemsaithong P, Miranda J, Chaiworapongsa T, Ravel J. 2014. The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. *Microbiome* **2**:4–4 DOI 10.1186/2049-2618-2-4.

Roxby AC, Fredricks DN, Odem-Davis K, Ásbjörnsdóttir K, Masese L, Fiedler TL, De Rosa S, Jaoko W, Kiarie JN, Overbaugh J, McClelland RS. 2016. Changes in vaginal microbiota and immune mediators in HIV-1-seronegative Kenyan women initiating depot medroxyprogesterone acetate. *Journal of Acquired Immune Deficiency Syndromes* **71**:359–366 DOI 10.1097/QAI.0000000000000866.

RStudio Team. 2016. RStudio: integrated Development for R. Boston: RStudio, Inc. *Available at http://www.rstudio.com/*.

Saunders S, Bocking A, Challis J, Reid G. 2007. Effect of *Lactobacillus* challenge on *Gardnerella vaginalis* biofilms. *Colloids and Surfaces B: Biointerfaces* **55**:138–142 DOI 10.1016/j.colsurfb.2006.11.040.

Smith BC, McAndrew T, Chen Z, Harari A, Barris DM, Viswanathan S, Rodriguez AC, Castle P, Herrero R, Schiffman M, Burk RD. 2012. The cervical microbiome over 7 years and a comparison of methodologies for its characterization. *PLOS ONE* **7**(**7**):e40425 DOI 10.1371/journal.pone.0040425.

Sobel JD. 2000. Bacterial vaginosis. *Annual Review of Medicine* **51**:349–356 DOI 10.1146/annurev.med.51.1.349.

Srinivasan S, Hoffman NG, Morgan MT, Matsen FA, Fiedler TL, Hall RW, Ross FJ, McCoy CO, Bumgarner R, Marrazzo JM, Fredricks DN. 2012. Bacterial communities in women with bacterial vaginosis: high resolution phylogenetic analyses reveal relationships of microbiota to clinical criteria. *PLOS ONE* **7**(**6**):e37818 DOI 10.1371/journal.pone.0037818.

Srinivasan S, Liu C, Mitchell CM, Fiedler TL, Thomas KK, Agnew KJ, Marrazzo JM, Fredricks DN. 2010. Temporal variability of human vaginal bacteria and relationship with bacterial vaginosis. *PLOS ONE* **5**(**4**):e10197 DOI 10.1371/journal.pone.0010197.

Srinivasan S, Munch MM, Sizova MV, Fiedler TL, Kohler CM, Hoffman NG, Liu C, Agnew KJ, Marrazzo JM, Epstein SS, Fredricks DN. 2016. More easily cultivated than identified: classical isolation with molecular identification of vaginal bacteria. *Journal of Infectious Diseases* **214**(**Suppl 1**):S21–S28 DOI 10.1093/infdis/jiw192.

Swidsinski A, Doerffel Y, Loening-Baucke V, Swidsinski S, Verstraelen H, Vaneechoutte M, Lemm V, Schilling J, Mendling W. 2010. *Gardnerella* biofilm involves females and males and is transmitted sexually. *Gynecologic and Obstetric Investigation* **70**:256–263 DOI 10.1159/000314015.

Tishkoff SA, Williams SM. 2002. Genetic analysis of African populations: human evolution and complex disease. *Nature Reviews Genetics* **3**:611–621 DOI 10.1038/nrg865.

Tokyol C, Aktepe OC, Cevrioglu AS, Altindis M, Dilek FH. 2004. Bacterial vaginosis: comparison of Pap smear and microbiological test results. *Modern Pathology* **17**:857–860 DOI 10.1038/modpathol.3800132.

Verstraelen H, Verhelst R, Claeys G, De Backer E, Temmerman M, Vaneechoutte M. 2009. Longitudinal analysis of the vaginal microflora in pregnancy suggests that *L. crispatus* promotes the stability of the normal vaginal microflora and that *L. gasseri* and/or *L. iners* are more conducive to the occurrence of abnormal vaginal microflora. *BMC Microbiology* **9**:116–116 DOI 10.1186/1471-2180-9-116.

Vodstrcil LA, Hocking JS, Law M, Walker S, Tabrizi SN, Fairley CK, Bradshaw CS. 2013. Hormonal contraception is associated with a reduced risk of bacterial vaginosis: a systematic review and meta-analysis. *PLOS ONE* **8(9)**:e73055 DOI 10.1371/journal.pone.0073055.

Vodstrcil LA, Twin J, Garland SM, Fairley CK, Hocking JS, Law MG, Plummer EL, Fethers KA, Chow EP, Tabrizi SN, Bradshaw CS. 2017. The influence of sexual activity on the vaginal microbiota and *Gardnerella vaginalis* clade diversity in young women. *PLOS ONE* **12(2)**:e0171856 DOI 10.1371/journal.pone.0171856.

Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73**:5261–5267 DOI 10.1128/AEM.00062-07.

Watts DH, Fazzari M, Fazarri M, Minkoff H, Hillier SL, Sha B, Glesby M, Levine AM, Burk R, Palefsky JM, Moxley M, Ahdieh-Grant L, Strickler HD. 2005. 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. *Journal of Infectious Diseases* **191**:1129–1139 DOI 10.1086/427777.

Wessels JM, Lajoie J, Vitali D, Omollo K, Kimani J, Oyugi J, Cheruiyot J, Kimani M, Mungai JN, Akolo M, Stearns JC, Surette MG, Fowke KR, Kaushic C. 2017. Association of high-risk sexual behaviour with diversity of the vaginal microbiota and abundance of *Lactobacillus*. *PLOS ONE* **12(11)**:e0187612 DOI 10.1371/journal.pone.0187612.

Van de Wijgert JH, Verwijs MC, Turner AN, Morrison CS. 2013. Hormonal contraception decreases bacterial vaginosis but oral contraception may increase candidiasis: implications for HIV transmission. *AIDS* **27**:2141–2153 DOI 10.1097/QAD.0b013e32836290b6.

Van Houdt R, Ma B, Bruisten SM, Speksnijder A, Ravel J, De Vries HJC. 2018. *Lactobacillus iners*-dominated vaginal microbiota is associated with increased susceptibility to *Chlamydia trachomatis* infection in Dutch women: a case-control study. *Sex Transm Infect* **94**:117–123 DOI 10.1136/sextrans-2017-053133.

Yang L, Hao Y, Hu J, Kelly D, Li H, Brown S, Tasker C, Roche NE, Chang TL, Pei Z. 2019. Differential effects of depot medroxyprogesterone acetate administration on vaginal microbiome in Hispanic White and Black women. *Emerging Microbes & Infections* **8**:197–210 DOI 10.1080/22221751.2018.1563458.

Zelezniak A, Andrejev S, Ponomarova O, Mende DR, Bork P, Patil KR. 2015. Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proceedings of the National Academy of Sciences of the United States of America* **112**:6449–6454 DOI 10.1073/pnas.1421834112.

Zhao S, Guo Y, Sheng Q, Shyr Y. 2014. Heatmap3: an improved heatmap package with more powerful and convenient features. *BMC Bioinformatics* **15**:P16 DOI 10.1186/1471-2105-15-S10-P16.

Zhou X, Brown CJ, Abdo Z, Davis CC, Hansmann MA, Joyce P, Foster JA, Forney LJ. 2007. Differences in the composition of vaginal microbial communities found in healthy Caucasian and Black women. *ISME Journal* **1**:121–133 DOI 10.1038/ismej.2007.12.

Zozaya-Hinchliffe M, Lillis R, Martin DH, Ferris MJ. 2010. Quantitative PCR assessments of bacterial species in women with and without bacterial vaginosis. *Journal of Clinical Microbiology* **48**:1812–1819 DOI 10.1128/JCM.00851-09.