The Vaginal Microbiome: The Influence of Intramuscular Depot-medroxyprogesterone Acetate Initiation on Vaginal Microbiota and A Comparison of PCR Approaches for Use in Predicting HIV Acquisition

Bridget Morrissey Whitney

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

Doctor of Philosophy

University of Washington

2020

Reading Committee:
Brandon L. Guthrie, Chair
Alison C. Roxby
Raymond Scott McClelland

Program Authorized to Offer Degree:
Epidemiology

University of Washington

Abstract

The Vaginal Microbiome: The Influence of Intramuscular Depot-medroxyprogesterone Acetate Initiation on Vaginal Microbiota and A Comparison of PCR Approaches for Use in Predicting HIV Acquisition

Bridget Morrissey Whitney

Chair of the Supervisory Committee:
Dr. Brandon L. Guthrie
Assistant Professor, Global Health
Assistant Professor, Epidemiology

The vaginal microbiome is a key factor in women's reproductive health and hormones may play an important role in the composition of vaginal bacterial communities. The vaginal microbiome is commonly evaluated using broad-range polymerase chain reaction (PCR) with next generation sequencing (NGS) or taxon-specific quantitative PCR (qPCR); both approaches have advantages and disadvantages. We sought to describe how DMPA-IM initiation influences the vaginal microbiome in postpartum women and provide evidence as to whether both PCR approaches provide the same value for predicting important biological outcomes of interest,

such as HIV acquisition, to help identify optimal PCR approach(s) for use in future research and clinical applications.

To assess the impact of DMPA-IM initiation on the vaginal microbiome, we used data from the Postpartum Family Planning (PFP) cohort, a prospective study of postpartum women in Kenya initiating DMPA-IM or non-hormonal contraception (non-HC). Enrollment occurred at the time of contraception initiation, approximately 6-14 weeks postpartum, and vaginal swabs for Nugent score determination, taxon-specific qPCR, and broad-range 16S rRNA gene PCR coupled with NGS were collected at enrollment and three-months post-initiation. Generalized estimating equations and linear mixed models were used to estimate mean change in Nugent score, total bacterial load, taxa concentrations, and bacterial community alpha diversity in women using DMPA-IM compared to those using non-HC. The effect of DMPA-IM on microbial community composition was assessed by comparing change in community type (CT) membership, created through hierarchical clustering, using a Wilcoxon signed-rank test.

The PFP cohort enrolled 54 HIV-negative women, 33 choosing DMPA-IM and 21 choosing non-HC. Women who chose DMPA-IM were more likely to be married (97% vs. 67%) and have resumed intercourse since delivery (52% vs. 29%) compared to women who chose non-HC. Baseline distributions of CTs was similar between contraceptive groups, however bacterial vaginosis by Nugent score was more common among DMPA-IM users. After three months, non-HC users were more likely to have *Lactobacillus* spp. dominant microbiomes compared to DMPA-IM users (n=11, 61% vs. n=8, 31%; χ^2 p=0.046). Nugent score decreased significantly among DMPA-IM users ($\triangle$=-1.89 points, p=0.02), however there was not a corresponding decrease in alpha diversity ($\triangle$=0.03, p=0.83). Among non-HC users, Nugent score remained more stable (△=-0.73 points, p=0.33) while alpha diversity decreased (△=-0.34, p=0.05). While there were significant changes in Nugent score and alpha diversity within contraceptive groups, the observed changes were not significantly different between the groups. After three months, significant decreases in the concentrations of *Sneathia* species, *Mycoplasma hominis*, and *Parvimonas* species Type 1 were seen among non-HC users, however concentrations remained stable among DMPA-IM users; contraceptive method was associated with significantly different changes in *M. hominis* concentration between groups (p=0.010).

To assess how abundance measures from qPCR and NGS compare in their relative value for predicting HIV acquisition, we used data from a case-control study evaluating the association between vaginal bacteria and HIV acquisition comprised of women from Eastern and Southern Africa at risk of HIV. A subset of 55 cases and 55 matched controls had both NGS and qPCR sequencing performed on their samples. We generated a series of models using logistic regression to assess the performance of bacterial concentration, from 9 selected taxa, and relative abundance measures as predictors of HIV. We performed principal components analysis (PCA) to assess the combined predictive value of bacterial abundances of taxa of interest. We also performed least absolute shrinkage and selection operator (LASSO) to select the most informative taxa for HIV prediction. Level of discrimination was quantified using the area under the receiver-operating curve (AUC) and AUCs were compared between models.

Within the nested case-control study, median concentrations and relative abundances of bacteria associated with HIV acquisition were low. A reference model containing demographic and risk factor variables had an acceptable discrimination value (AUC=0.706, 95% Confidence Interval (CI): 0.609-0.804). The first three principal components (PCs) from qPCR data resulted in an AUC=0.787 (95%CI: 0.702-0.872) and the first three PCs from NGS data resulted in an AUC=0.761 (95%CI: 0.672-0.850). AUCs from LASSO selection resulted in similar discrimination (AUC=0.771 vs. AUC=0.784, respectively); there was no significant difference in discrimination between PCR approaches.

In a cohort of postpartum African women, vaginal bacterial diversity did not change in DMPA-IM users despite a reduction in Nugent-BV, but decreased significantly among women using non-HC. Choice of contraception may influence Lactobacillus recovery in postpartum women. Additionally, our findings suggest that postpartum use of DMPA-IM and non-HC may have differential impacts on the vaginal concentrations of some bacteria that have previously been associated with HIV acquisition. In the nested case-control study, we found the abundance of bacterial taxa associated with HIV acquisition from qPCR, in the form of concentration, and from broad-range PCR with NGS, in the form of relative abundance, provide similar predictive discrimination between women who acquired HIV and women who remained HIV uninfected.

TABLE OF CONTENTS

LIST OF TABLES

Table 4.6. HIV discrimination value from models selected using least absolute shrinkage and selection operator (LASSO).

LIST OF SUPPLEMENTAL TABLES

LIST OF FIGURES

LIST OF SUPPLEMENTAL FIGURES

Figure S3.4. (A) Nugent score, (B) total bacterial load, and (C) concentration of vaginal taxa over time among DMPA-IM users only with fitted trend lines for mean change among women ≥1 detectable value during follow-up.

Figure S4.1. Receiver-operating curves for HIV discrimination from: (A,B) detection models using the first three principal components from principal components analysis (PCA) and (C,D) detection models selected using least absolute shrinkage and selection operator (LASSO).

CHAPTER 1. Introduction

Modern contraceptives and intramuscular depot-medroxyprogesterone acetate

Modern contraceptives promote maternal and child health and reduce unintended pregnancies, abortions, and pregnancy-related deaths.[1] The most recent estimates (2017) indicate that modern contraceptive use prevents approximately 308 million unintended pregnancies every year in low- and middle-income countries, and if unmet need for modern contraception were satisfied it is estimated that there would be a three-quarters reduction in unintended pregnancies (from 89 million to 22 million), unplanned births (from 30 million to 7 million), and induced abortions (from 48 million to 12 million) annually, as well as 76,000 fewer maternal death per year.[2]

Injectable hormonal contraceptives (IHCs), including intramuscular depot-medroxyprogesterone acetate (DMPA-IM), are the most commonly used contraceptives among women in sub-Saharan Africa, where the burden of vaginal dysbiosis is highest [3] and 25.7 million people live with HIV. [4] IHC use has risen dramatically in the last two decades and now accounts for approximately 40% of all contraceptives used in Eastern and Southern Africa.[5] DMPA-IM, a progestin-only contraceptive, inhibits the secretion of pituitary gonadotropins, resulting in anovulation and decreased production of estrogen.[6] DMPA-IM is highly effective and suitable for use in breastfeeding women more than weeks postpartum because progestins do not suppress production of lactation.[7,8] Additionally, DMPA-IM is administered every 13 weeks,[1,9] making it a discreet and convenient form of birth control for women, as well as an efficient form of contraception to deliver from a health systems standpoint.[10]

The vaginal microbiome

The vaginal microbiome plays a key role in women's reproductive health. Hydrogen peroxide-producing *Lactobacillus*-dominant vaginal bacterial communities are considered optimal for health;[11,12] optimal microbiota are associated with favorable sexual and reproductive health outcomes and a lack of vaginal symptoms and/or genital inflammation.[13] Communities dominated by anaerobic bacteria, or non-optimal microbiota, as well as specific microbial taxa, are associated with subclinical inflammation, poor reproductive health outcomes, and sexually transmitted infections.[14-17] Bacterial vaginosis measured by Nugent score (Nugent-BV), the most thoroughly studied nonoptimal vaginal microbiota,[3] has been associated with a 60% increase in the relative risk of acquiring HIV in women,[18] and treatment of Nugent-BV has resulted in lower incidence of some bacterial STIs.[19,20] Culture-independent molecular methods have improved our understanding of the composition of the vaginal microbiota, but the factors shaping vaginal bacterial communities are still not well understood.[21]

Hormones, including estrogens and progestogens, play important roles in vaginal microbial ecology,[22-24] and exogenous hormones, such as contraceptives, may induce important changes in the composition of vaginal microbiota and production of soluble factors by bacteria. Reduced levels of estrogen have been associated with glycogen suppression,[23,25] and glycogen is an important substrate for *Lactobacillus* species.[25-27] Consequently, a reduction of glycogen in the female reproductive tract could shift the vaginal microbiome to an anaerobic, non-optimal state. Changes in vaginal bleeding patterns and perturbation of the regular menstrual cycle associated with DMPA-IM initiation may also impact the vaginal microbiome.[24,28]

DMPA-IM and HIV

Numerous observational studies have reported an association between DMPA-IM use and increased HIV acquisition risk, with an average estimate of a 40-50% higher hazard of HIV acquisition among DMPA-IM users compared to non-hormonal contraceptive (non-HC) users.[29] Several biologically plausible mechanisms to explain the increased risk of HIV acquisition among DMPA-IM users were hypothesized, including, but not limited to: 1. Changes in the female genital tract, including vaginal thinning; 2. Changes in local and/or systemic inflammation; 3. Increased proliferation of HIV in target cells; and 4. Changes in the composition of vaginal microbiota and production of soluble factors by bacteria.[23,30] Recently released results from The Evidence for Contraceptive Options and HIV Outcomes (ECHO) Trial,[31] which did not find differences in HIV acquisition across the contraceptive methods evaluated, have greatly reduced clinical concern about DMPA-IM use in women at risk for HIV; however, there is still uncertainty in the field as to whether DMPA-IM increases the risk of HIV acquisition relative to non-HC methods.[32] Additionally, it is important to understand if and how DMPA-IM affects the vaginal environment for other women's health outcomes.

DMPA-IM and the vaginal microbiome

A limited number of studies have evaluated DMPA-IM's effect on the vaginal microbiota; most previous studies were cross-sectional and not all were specifically designed to assess the impact of DMPA-IM initiation on vaginal bacteria. Additionally, there is a gap in this data for postpartum women. Data generally indicate that DMPA-IM reduces Nugent-BV[33-36] and colonization by BV-associated bacteria,[37] although null associations between DMPA-IM and Nugent-BV have also been reported.[38-40] Of studies reporting culture-independent molecular data, results are heterogenous.[37-43] Longitudinal studies have reported that DMPA-IM use is associated with: (1) higher levels of *G. vaginalis*,[39] (2) no significant changes in concentrations of *G. vaginalis*, *A. vaginae*, and *Megasphaera* phylotype 1,[38] and (3) marked decreases in the concentration of *G. vaginalis*.[40] Additionally, DMPA-IM use has been associated with no change[38] or decreases[25,39,42] in *Lactobacillus* colonization. Studies assessing DMPA-IM's effect on overall vaginal microbial community composition have been similarly heterogeneous.[41,42,44] Social, geographic, and cultural factors may play a role in these observed differences.

Evaluation of the vaginal microbiome

The vaginal microbiome is commonly evaluated using broad-range polymerase chain reaction (PCR) with next generation sequencing (NGS), hereafter referred to as NGS, or taxon-specific quantitative PCR (qPCR). Both approaches have advantages and disadvantages. NGS targets conserved regions of the 16S rRNA gene and is advantageous because it provides a comprehensive picture of the microbiome in a sample and does not require the *a priori* determination of which bacterial taxa will be assessed. Advances in NGS have allowed sampling of bacterial DNA at new depths, providing a more detailed picture of the vaginal microbiome.[45] NGS is useful for assessing community composition and bacterial diversity, but it only provides the relative abundance of bacterial taxa present in a specimen.[45] Additionally, analytic decisions at all stages of processing, including the operational taxonomic unit (OTU) generation strategy, choice of reference database, clustering algorithm, and software package, can greatly affect results.[46,47] In contrast, qPCR measures the absolute concentration of pre-defined bacterial taxa through the use of taxon-specific primers. Data describing taxon-specific concentrations enables the evaluation of changes in bacterial concentration over time, which may be inapparent from relative abundance measures. However, qPCR does not provide an overall view of vaginal microbiota, the major limitation of this approach.[45] Using both PCR approaches together provides complementary data. To our knowledge, no published studies have directly compared the predictive value of different measures of the vaginal microbiota for biological outcomes such as HIV acquisition.

Data for this dissertation come from four cohorts in Eastern and Southern Africa. The first two chapters utilize data from the Postpartum Family Planning (PFP) cohort, a prospective cohort study of DMPA-IM initiation in postpartum women in Nairobi, Kenya. Data for the third chapter come from a nested case-control study, evaluating the association between specific vaginal bacteria and risk of HIV acquisition, including data from The Mombasa Cohort, The Mama Salama Study, and The Partners PrEP Study. We begin by presenting data on the effect of DMPA-IM initiation in the postpartum period on vaginal microbial community composition and diversity (Chapter 2). Next, we describe the effect of DMPA-IM initiation on eight bacterial taxa previously associated with HIV acquisition (Chapter 3). Finally, we present results from predictive model study comparing how abundance measures from qPCR and NGS compare and their relative value for predicting HIV (Chapter 4). We conclude by summarizing the findings of these studies and discussing their implications (Chapter 5).

CHAPTER 2: Influence of postpartum intramuscular depot-medroxyprogesterone acetate initiation on vaginal microbiota

ABSTRACT

Background: The vaginal microbiome is a key factor in women's reproductive health and hormones may play an important role in the composition of vaginal bacterial communities.

Methods: Vaginal swabs were collected from a cohort of postpartum Kenyan women initiating intramuscular depot-medroxyprogesterone acetate (DMPA-IM) or non-hormonal contraception (non-HC). Women were evaluated for bacterial vaginosis by Nugent score (Nugent-BV). Vaginal bacterial community composition was assessed using broad-range 16S rRNA gene PCR with high-throughput sequencing and hierarchical clustering was used to create vaginal community types (CTs). The effect of DMPA-IM on microbial community composition was assessed by comparing change in CT membership using a Wilcoxon signed-rank test. Changes in Nugent score, total bacterial load, and Shannon's diversity index (SDI) from enrollment to three-months post-initiation were estimated using multivariable linear mixed effects regression.

Results: Among 54 HIV-negative women, 33 choosing DMPA-IM and 21 choosing non-HC, baseline distributions of bacterial community types were similar, however Nugent-BV was more common among DMPA-IM users. After three months, non-HC users were more likely to have *Lactobacillus* spp. dominant microbiomes compared to DMPA-IM users (n=11, 61% vs. n=8, 31%; χ^2 p=0.046). Nugent score decreased significantly among DMPA-IM users (△=-1.89 points, p=0.02), however there was not a corresponding decrease in SDI (△=0.03, p=0.83). Among non-HC users, Nugent score remained more stable (△=-0.73 points, p=0.33) while SDI decreased (△=-0.34, p=0.05). While there were significant changes in Nugent score and alpha diversity within contraceptive groups, the observed changes were not significantly different between the groups. Total bacterial load, which was similar at enrollment, increased slightly in non-HC users and decreased slightly in DMPA-IM users, making the difference in change of total bacterial load between the two contraceptive groups significant (difference=-0.64 [95%CI: -1.19 to -0.08], p=0.02).

Conclusions: Postpartum vaginal bacterial diversity did not change in DMPA-IM users despite a reduction in Nugent-BV, but decreased significantly among women using non-HC. Choice of contraception may influence *Lactobacillus* recovery in postpartum women.

INTRODUCTION

The vaginal microbiome plays a key role in women's reproductive health. It is widely accepted that vaginal microbiota dominated by *Lactobacillus* species, particularly *Lactobacillus crispatus*, are optimal and promote vaginal health;[3,11-13,21,48] optimal microbiota are associated with favorable sexual and reproductive health outcomes and a lack of vaginal symptoms and/or genital inflammation.[13] Conversely, vaginal microbiota deplete of *Lactobacillus* spp. with high abundances of anaerobic bacteria, or nonoptimal microbiota, are associated with subclinical inflammation, poor reproductive outcomes, pelvic inflammatory disease (PID), and increased incidence of sexually transmitted infections (STIs).[12,14,15,48-50] Bacterial vaginosis measured by Nugent score (Nugent-BV), the most thoroughly studied nonoptimal vaginal microbiota,[3] has been associated with a 60% increase in the relative risk of acquiring HIV in women,[18] and treatment of Nugent-BV has resulted in lower incidence of some bacterial STIs.[19,20] Culture-independent molecular methods have improved our understanding of the composition of the vaginal microbiota, but the factors shaping vaginal bacterial communities are still not well understood.[21]

Endogenous and exogenous sex hormones, including estrogens and progestogens, are known to influence vaginal bacterial communities;[22-24,36] consistent evidence suggests that high levels of estradiol promote *Lactobacillus* spp., especially *L. crispatus*.[21] Intramuscular depot-medroxyprogesterone acetate (DMPA-IM), a progestin-only injectable contraceptive, induces hypoestrogenism,[25,51,52] which could potentially shift the vaginal microbiota toward an anaerobic, polymicrobial composition if hypoestrogenism leads to loss of vaginal glycogen needed for growth of lactobacilli.[53]

Numerous observational studies have reported an association between DMPA-IM use and increased HIV acquisition risk, with an average estimate of a 40-50% higher hazard of HIV acquisition among DMPA-IM users compared to non-hormonal contraceptive (non-HC) users.[29] Recently released results from The Evidence for Contraceptive Options and HIV Outcomes (ECHO) Trial,[31] which did not find differences in HIV acquisition across the contraceptive methods evaluated, have greatly reduced clinical concern about DMPA-IM use in women at risk for HIV; however, there is still uncertainty in the field as to whether DMPA-IM increases the risk of HIV acquisition relative to non-HC methods.[32] Additionally, it is important to understand if and how DMPA-IM affects the vaginal environment for other women's health outcomes. A limited number of studies have evaluated DMPA-IM's effect on the vaginal microbiota; most previous studies were cross-sectional and not all were specifically designed to assess the impact of DMPA-IM initiation on vaginal bacteria. Data generally indicate that DMPA-IM reduces Nugent-BV[33-36] and colonization by BV-associated bacteria,[37] although null associations between DMPA-IM and Nugent-BV have also been reported.[38-40] Of studies reporting culture-independent molecular data, results are heterogenous.[37-43] Longitudinal studies have reported that DMPA-IM use is associated with: (1) higher levels of *G. vaginalis*,[39] (2) no significant changes in concentrations of *G. vaginalis*, *A. vaginae*, and *Megasphaera* phylotype 1,[38] and (3) marked decreases in the concentration of *G. vaginalis*.[40] Additionally, DMPA-IM use has been associated with no change[38] or decreases[25,39,42] in *Lactobacillus* colonization. Studies assessing DMPA-IM's effect on overall vaginal microbial community composition have been similarly heterogeneous.[41,42,44] Social, geographic, and cultural factors may play a role in these observed differences.

We designed an observational cohort to explore the immediate effect of DMPA-IM initiation on vaginal microbial community composition, to understand if DMPA-IM use is associated with changes to vaginal microbiota, creating a potential pathway between DMPA-IM use and harmful sequelae, including STIs and PID. We hypothesized that women who initiate DMPA-IM would have increased vaginal taxa richness and diversity.

METHODS

Study setting, subjects, and design

We conducted a prospective study of postpartum women initiating contraception, designed specifically to assess how DMPA-IM initiation affects the vaginal microenvironment. From a primary care clinic in Nairobi, Kenya, we recruited women at 6-14 weeks postpartum and interested in initiating DMPA-IM or a non-HC method, including condoms, lactational amenorrhea, or rhythm. Women were ineligible if already using HC or unwilling to learn their HIV status, and excluded if there was evidence of cervicitis or STI.

Women chose their preferred contraceptive method at enrollment; participants who chose DMPA-IM received their first injection from study staff following collection of study specimens, enabling documentation of exact timing and duration of DMPA-IM exposure. Women using DMPA-IM were asked to return after 9-14 days for a visit (hereafter designated the two-week post-injection visit) at time of pharmacokinetic peak plasma medroxyprogesterone acetate (MPA) concentration.[51,52] All women were asked to return approximately three months post-enrollment for a follow-up visit (hereafter designated the three-month follow-up visit). Questionnaires on health and sexual activity were completed at all visits. Vaginal swabs for microbiota evaluation were collected during pelvic exams; swabs were taken bilaterally in the space that includes right and left lateral fornix and the distal two-thirds of the vaginal wall. Swabs were not collected from women with vaginal bleeding, including spotting; women were asked to return for swab collection when bleeding had stopped. Vaginal swabs were cryopreserved at –80°C and sent to the Fred Hutchinson Cancer Research Center (Seattle, WA, USA) for analysis.

Laboratory procedures

Peripheral blood was used for point-of-care HIV testing (Determine rapid enzyme-linked immunosorbent assay, Abbott, Abbott Park, Illinois). Serum MPA concentration was determined using a quantitative liquid chromatography mass spectrometry assay (Biomarkers Core Laboratory, Columbia University).[54] Vaginal fluid was evaluated by Gram stain by a highly trained technician for Nugent score determination; a score of ≥7 was classified as Nugent-BV.[55] Slides were read within one week of collection and the technician was blinded to any identifying details.

DNA extraction, quantification, and Polymerase Chain Reaction (PCR)

DNA was extracted using the BiOstic Bacteremia DNA Isolation Kit (Mobio, Carlsbad, CA, now Qiagen). To assess contamination, sham digests from swabs without human contact were processed; no-template water controls were included with every PCR assay. DNA samples were tested for PCR inhibitors,[56] and 16S rRNA gene-copies in each sample (total bacterial load) were measured using a broad-range qPCR assay.[57] Extracted DNA was subjected to broad-range PCR targeting the V3-V4 region of the 16S rRNA gene with next-generation sequencing by Illumina MiSeq (Illumina, San Diego, CA).[58] Sequence reads were filtered for length, quality score, and contaminants, and classified using the phylogenetic placement tool, *pplacer*, and a reference set of vaginal bacteria.[59] Laboratory staff were blinded to contraception status of participants.

Bacterial community analyses

Sequence read counts were converted to proportions to determine the relative abundance of each taxa. Alpha diversity was quantified using Shannon's diversity index (SDI), a measure of taxa richness and evenness.[60] Beta diversity between samples was calculated using weighted UniFrac distances[60] and visualized using principal coordinates analysis (PCoA). Hierarchical clustering based on bacterial composition and abundance, as well as weighted UniFrac distances, was conducted using Unweighted Pair Group Method with Arithmetic Mean (UPGMA) and used to create vaginal community types (CTs);[61] CTs with five or less observations were pooled together as "other." Linear discriminant analysis with effect size (LEfSe) was used to identify differentially abundant taxa between DMPA-IM users and non-HC users cross-sectionally at enrollment and follow-up using default parameters of α=0.05 and a minimum linear discriminant analysis (LDA) score of 2.0.[62]

The effect of DMPA-IM on microbial community composition was assessed by comparing change in CT membership among the 44 women who returned for the three-month follow-up visit using a Wilcoxon signed-rank test. To visualize similarities in the overall variation in data between the timepoints, the enrollment visit's first PCoA vector (PCoA1) was plotted against the three-month follow-up visit's PCoA1, distinguished by contraceptive group, and Pearson correlation coefficients between the two visits were calculated. The effect of DMPA-IM on Nugent score, total bacterial load, and SDI was estimated using multivariable linear mixed effects regression with an interaction term between contraceptive group and time (days from enrollment to swab collection). Models were adjusted for days from delivery to enrollment, as well as potential a priori confounders, including age, marital status, and resumption of intercourse post-delivery.

Within the DMPA-IM group only, we used similar methods to evaluate changes in outcomes from enrollment to the two-week post-injection visit, when MPA serum concentrations peak,[51,52] and from the two-week post-injection visit to the three-month follow-up visit. We also evaluated the potential for a dose-response relationship between peak serum MPA concentration and Nugent score, total bacterial load, and SDI at the two-week post-injection and the three-month follow-up visits using multivariable linear regression. Serum MPA concentrations were categorized into tertiles for analysis: (1) 0.69-1.09 ng/ml, (reference category), (2) 1.10-1.50 ng/ml, and (3) 1.51-3.54 ng/ml. Statistical analyses were conducted using R version 3.6.1, the Galaxy web platform,[63] and Stata version 15.

The study was approved by the Kenyatta National Hospital Ethics and Research Committee and the University of Washington Institutional Review Board. Written informed consent was obtained in English or Kiswahili from all participants.

RESULTS

Participant characteristics

We enrolled 54 HIV-negative women. Mean age was 24.2 years (standard deviation (SD)=4.6) and a majority of women (n=35, 65%) had less than a secondary school education (Table 2.1). More than half of the women enrolled chose DMPA-IM (N=33, 61%). A higher proportion of women who chose DMPA-IM were married (n=32, 97% vs. n=14, 67%) and had resumed sexual intercourse post-delivery (n=17, 52% vs. n=6, 29%). Women who chose DMPA-IM started contraception sooner after delivery compared to women who chose non-HC methods (7.1 weeks vs. 9.9 weeks). At enrollment, Nugent-BV was more prevalent in the DMPA-IM group compared to the non-HC group (n=19, 58% vs. n=7, 33%), but not statistically significantly different (p>0.05). Of enrolled women, 44 (81%) returned for three-month follow-up, with a slightly greater proportion of non-HC users retained in the cohort (n=18, 86% of non-HC users vs. n=26, 79% of DMPA-IM users). Time between the enrollment visit and the three-month follow-up visit was highly comparable between the contraceptive groups (84 days for non-HC users [interquartile range (IRQ): 84-90 days] and 84 days for DMPA-IM users [IRQ: 84-89 days]).

Vaginal bacterial communities among women using DMPA-IM and women using non-HC

Vaginal bacterial community composition among all women at both enrollment and the three-month follow-up visit (N=98 swabs from 54 women) and metadata superposed on the clustering tree are shown in Figure 2.1. *Lactobacillus iners* was the only commonly detected *Lactobacillus* spp. Women with diverse, heterogeneous communities, corresponding to higher SDI values, were more likely to have an intermediate or BV-positive Nugent score compared to women colonized primarily with *L. iners* (χ^2 p<0.001).

Five bacterial community types (CTs) were identified using UPGMA and combining small clusters with high diversity. Two CTs were dominated (mean relative abundance ≥75% of total composition) by a single taxon, CT1 by *L. iners* (n=37, 38%) and CT2 by *Gardnerella vaginalis* (n=15, 15%), and the remaining three CTs contained multiple bacterial taxa without a dominant bacterium (Table 2.2). CT3 (n=9, 9%) had substantial (>10%) abundances of *G. vaginalis* and

Lactobacillus spp., and CT4 (n=21, 21%) had substantial abundances of *G. vaginalis*, *Prevotella* spp., and *A. vaginae*. CT5 (n=16, 16%) was a heterogenous mix of anaerobic bacteria. PCoA confirmed clustering of CTs, with the poorest separation between CT4 and CT5 (Figure S2.1). The first three principal coordinates explained approximately 84% of the variation in the data (PCoA1: 57%; PCoA2: 21%; PCoA3: 6%). Assessment of clustering of CTs by visit via PCoA did not show appreciable differences between time points (Figure S2.2).

CT1 was the least diverse cluster (mean SDI=0.24 [SD=0.29]) and CT5 was the most diverse cluster (mean SDI=2.12 [SD=0.85]), with the interim CTs each significantly more diverse than the previous, except for CT4 and CT5 (Table 2.2). At enrollment, CT membership and SDI did not differ significantly between contraceptive groups, although women who chose DMPA-IM had a lower proportion of women in the *Lactobacillus* spp. dominant CT1 compared to the women who chose non-HC (n=10, 30% vs. n=8, 38%, respectively) and a higher proportion of women in the *G. vaginalis* dominant CT2 (n=7, 21% vs. n=2, 10%, respectively), aligning with the observation that more women in the DMPA-IM group had Nugent-BV at enrollment (Table 2.3). CT membership and SDI did not differ significantly between contraceptive groups at the three-month follow-up visit. To conserve power to detect differences between contraceptive groups, we grouped CTs by *Lactobacillus* spp. dominant, or more optimal microbiota, (CT1) versus *Lactobacillus* spp. deplete CTs, or less optimal microbiota, (CTs 2-5); we found that non-HC users were more likely to have *Lactobacillus* spp. dominant microbiomes at the three-month follow-up visit compared to DMPA-IM users (n=11, 61% vs. n=8, 31%, respectively; χ^2 p=0.046).

LEfSe analyses detected differences in the relative abundance of several taxa associated with nonoptimal vaginal communities between DMPA-IM users and non-HC users at both the enrollment and follow-up visit (Figure S2.3). At enrollment, *G. vaginalis* and two *Porphyromonas* species were more abundant in DMPA-IM users compared to non-HC users, and *Mycoplasma hominis* and two other anaerobic taxa were more abundant in non-HC users compared to DMPA-IM users. At the three-month follow-up visit, several anaerobic taxa, including *Anaerococcus* species, *Dialister propionicifaciens*, *Finegoldia magna*, *Peptoniphilus* species, and *Porphyromonas* species, were significantly more abundant in DMPA-IM users compared to non-HC users, whereas only one taxon, *Prevotella disiens*, was more abundant in non-HC users compared to DMPA-IM users.

Patterns of change among DMPA-IM users compared to non-HC users

While there was no significant difference in CT membership between contraceptive groups at follow-up, given the significant difference in *Lactobacillus* spp. dominant and *Lactobacillus* spp. deplete CTs, we explored qualitative patterns of change in CT membership among the 44 women who returned for the three-month follow-up visit. Non-HC users appeared to shift from *Lactobacillus* spp. deplete CTs to CT1 (*L. iners* dominant) more frequently than DMPA-IM users (17% vs. 4%, respectively) (Figure 2.2A). Additionally, zero non-HC users who started in CT1 shifted to a less optimal CT, while two women using DMPA-IM shifted from CT1 to *Lactobacillus* spp. deplete CTs. Movement in and out of CT4, the CT with the greatest prevalence of Nugent-BV, also appeared to be different: approximately a quarter of women in each contraceptive group started in CT4, but by the three-month follow-up visit only 11% of non-HC users were in CT4 whereas 24% of DMPA-IM users were grouped into CT4. In general, movement between *Lactobacillus* spp. deplete CTs was common, and more DMPA-IM users changed CTs during follow-up compared to non-HC users (50% vs. 33%). Comparison of the enrollment visit's first principal coordinate and the three-month follow-up visit's first principal coordinate showed high correlation by contraceptive group (non-HC: r=0.63, p<0.001; DMPA-IM: r=0.80, p<0.001) (Figure S2.4), indicating similar patterns in composition within contraceptive group at both visits. Assessment of clustering of CTs by visit via PCoA did not show appreciable differences between time points (data not shown).

Over three months, mean Nugent score decreased by 0.73 points among women using non-HC (95%CI: -2.18 to 0.73, p=0.33) and 1.89 points among women using DMPA-IM (95%CI: -3.53 to -0.25, p=0.02), making the groups more similar in Nugent-BV status at follow-up (Table 2.4, Figure 2.2B). Alpha diversity did not mirror this change: non-HC users had a significant decrease in diversity (△=-0.34 [95%CI: -0.67 to -0.001], p=0.05), while DMPA-IM users did not have any change in measured diversity at three-months post-injection (△=0.03 [95%CI: -0.24 to 0.30], p=0.83). While there were significant changes in Nugent score and alpha diversity within contraceptive groups, the observed changes were not significantly different between the groups (difference=-1.17 [95%CI: -3.35 to 1.02], p=0.30 and difference=0.37 [95%CI: -0.07 to 0.80], p=0.10, respectively). Total bacterial load, which was similar at enrollment (Table 2.3), increased slightly over time in non-HC users and decreased slightly over time in DMPA-IM users, making the difference in change of total bacterial load between the two contraceptive groups significant (difference=-0.64 [95%CI: -1.19 to -0.08], p=0.02) (Table 2.4).

Vaginal bacterial communities among DMPA-IM users only

Incorporating data from the peak MPA timepoint in the DMPA-IM group, vaginal bacterial community composition (N=88 swabs from 33 women) was similar to the pattern seen in all women described in Figure 2.1, with the exception that only four distinct CTs were identified using UPGMA (Figure S2.5, Table S2.1); CT3, a mixture of *G. vaginalis*, *Lactobacillus* spp., and other anaerobic bacteria, was not clearly separated from the other clusters in this subset of samples.

Patterns of change among DMPA-IM users only

There was no remarkable pattern in change in CT membership in the first two weeks post-injection among the 30 women who returned for at least one follow-up visit (Figure S2.6A,B); there was some movement between *Lactobacillus* spp. dominant and *Lactobacillus* spp. deplete CTs, but in small numbers and both directions, and movement between *Lactobacillus* spp. deplete CTs was common. There was also no remarkable pattern in change in Nugent-BV status in the first two weeks post-injection (Figure S2.6C).

Using data from just the enrollment and three-month follow-up visits, Nugent score, SDI, and total bacterial load all appeared to be stable, however our interim sampling identified time-dependent fluctuations in Nugent score and total bacterial load consistent with changing MPA kinetics. Mean Nugent score decreased more rapidly in the first two weeks post-injection (△=-1.31 [95%CI: -2.77 to 0.15], p=0.08) than in the period between the two-week follow-up visit and the three-month follow-up visit (△=-0.62 [95%CI: -2.15 to 0.91], p=0.43) (Figure 2.3). Similarly, total bacterial load decreased significantly in the first two weeks post-injection (△=-0.63 [95%CI: -0.98 to -0.28]; p<0.001), but mostly rebounded by the three-month follow-up visit (△=0.44 [95%CI: -0.05 to 0.95]; p=0.08), making total bacterial load appear relatively stable over follow-up. We did not identify a dose-response relationship between serum MPA concentration and any of the outcomes measured (Table S2.2).

DISCUSSION

In a cohort of African women initiating postpartum contraception, we observed several patterns that indicate that DMPA-IM use may be associated with the maintenance of nonoptimal, *Lactobacillus* deplete microbial communities, even when Nugent-BV is reduced. DMPA-IM was not associated with increases in nonoptimal bacterial communities in our cohort. This was demonstrated by a significant reduction in Nugent score but no corresponding change in alpha diversity or movement to *Lactobacillus* dominant microbial communities among DMPA-IM users. DMPA-IM was not associated with increases in nonoptimal bacterial communities in our cohort, as defined by increases in alpha diversity or movement from *Lactobacillus* dominant to *Lactobacillus* deplete communities; among DMPA-IM users, seven of the nine women who started in CT1 remained in CT1. Very few women had *L. crispatus* present at either time period, limiting our ability to assess movement between *L. crispatus* dominant communities, considered the most optimal composition, to intermediate, *L. iners*-dominated communities.

As the understanding of the vaginal microbiome has evolved, specific patterns among subsets of women have been identified. Recent studies have found that the vaginal microbiome during pregnancy and the postpartum period may be different from the vaginal microbiome found in non-pregnant women of reproductive age. Women in the immediate postpartum period have been observed to have more diverse vaginal microbiota, with lower levels of *Lactobacillus* spp., compared to both pre-pregnancy and during pregnancy.[64,65] DMPA-IM initiation in the postpartum period may maintain these nonoptimal vaginal microbial communities, characterized by depletion of *Lactobacillus* spp., by allowing a hypoestrogenic state to persist, while non-HC users return to a more optimal microbiota as the menstrual cycle returns. The hypoestrogenism induced by DMPA-IM has been associated with glycogen suppression,[25,66] and glycogen is an important substrate for *Lactobacillus* spp.[25,26] This hypothesis aligns with our findings: DMPA-IM initiation was not associated with large increases in diversity, however it appeared that DMPA-IM exposure might support continued dominance by non-*Lactobacillus* taxa compared to women using non-HC.

While we observed a significant decrease in Nugent score among women using DMPA-IM, which is congruent with some existing literature,[33-36] we did not see a change in bacterial community alpha diversity; even though Nugent-BV was reduced, women still had diverse, nonoptimal vaginal microbiomes. This finding is congruent with the concept that there are different types of nonoptimal vaginal bacterial communities beyond Nugent-BV.

Our finding of a reduction in Nugent score among women initiating DMPA-IM is different from recently published results from a cohort of HIV-negative, nonpregnant women in Zimbabwe,[38] which found that use of DMPA-IM did not significantly change the concentration of three BV-associated species, G*. vaginalis*, *Atopobium vaginae*, and *Megasphaera*-like bacterium, nor did it result in lower prevalence of BV. There are several possible explanations for the discrepancy in results related to DMPA-IM's association with BV, including that there was a much lower prevalence of Nugent-BV among DMPA-IM users at enrollment in the Zimbabwean cohort(29%) compared to our study (58%), the use of different populations (postpartum women vs. healthy, non-pregnant women with regular menstrual cycles), and the use of different parameterizations of BV (continuous change in Nugent score here vs. binary diagnosis of Nugent-BV in the Zimbabwean cohort). Additionally, in our study women who chose DMPA-IM had higher Nugent scores at enrollment compared to women who chose non-HC, giving them more latitude to decrease in Nugent score. The discrepancy in Nugent score at enrollment may be explained by a shorter time interval from delivery to contraceptive initiation and/or a higher proportion of DMPA-IM users reporting resumption of sexual intercourse at enrollment.

A major strength of this study is that participants were known to be contraceptive-free prior to enrollment and received their first DMPA-IM injection from study staff, confirming timing and duration of exposure; potential contamination from other contraceptives and uncertainty in type of injectable contraceptive are not of concern. Breastfeeding behaviors were uniform in both contraception groups and antibiotic use was very low in the cohort. The study design also allowed for comparison of pre- and post-DMPA-IM effects, as women could serve as their own controls. Another strength is that specimens were collected when serum MPA levels peak, allowing for examination of a dose-response relationship between MPA and the outcomes of interest.

The generalizability of our study may be limited due to the source population of participants; postpartum, amenorrheic women were specifically enrolled to reduce expected variability of the vaginal microbiome due to the menstrual cycle;[24,28,67] however, vaginal microbiota in the postpartum period appears to be different from vaginal microbiota found in menstruating women of childbearing age.[64,65] A second limitation is that participants selected their own contraceptive method, and women who rely on non-HC methods alone may be behaviorally and demographically different from women who choose DMPA-IM.[68] While we collected and adjusted for important confounders in our statistical models, the observed differences between groups could be due to residual confounding. Lastly, while our study was well-powered for continuous measures, we had limited power to assess changes in bacterial community type during follow-up, and had a modest loss to follow up (19% of women), further reducing our power to detect true differences.

Overall, these findings suggest that nonoptimal vaginal bacterial communities are common among Kenyan women, and that DMPA-IM use among postpartum women may be associated with maintenance of highly diverse, nonoptimal vaginal bacterial communities. While our cohort was small and restricted to a specific population of women, these findings suggest that DMPA-IM's effect on the vaginal microenvironment deserves further attention, especially in postpartum women. DMPA-IM is suitable for use in breastfeeding women four or more weeks postpartum because progestins do not suppress production of breast milk,[7] making it one of the few approved highly effective contraceptives in the postpartum period. Additionally, while the ECHO trial has relieved apprehension about use of DMPA-IM in women in high HIV prevalence settings,[31] nonoptimal vaginal microbiota are a risk factor for other negative health outcomes, including PID and the acquisition of other STIs; DMPA-IM use in the postpartum period may pose greater risks than it does in the non-postpartum period, due to high vaginal microbial diversity seen in this period and the potential for DMPA-IM to stabilize nonoptimal vaginal composition.

TABLES AND FIGURES

Table 2.1. Characteristics of postpartum women at enrollment (N=54).

Characteristic[a]	**All (n=54)**	**Non-hormonal contraception (n=21)**	**DMPA-IM (n=33)**
Age (years)	24.2 (4.6)	26.0 (6.0)	23.1 (3.1)
Married	46 (85.2%)	14 (66.7%)	32 (97.0%)
Education level at enrollment			
Less than secondary school	35 (64.8%)	13 (61.9%)	22 (66.7%)
Secondary school or more	19 (35.2%)	8 (38.1%)	11 (33.3%)
Lifetime number of partners	1.9 (1.1)	1.7 (1.0)	2.0 (1.1)
Number of pregnancies	1.9 (1.2)	2.1 (1.4)	1.8 (1.0)
Vaginal delivery (most recent delivery)	46 (85.2%)	18 (85.7%)	28 (84.8%)
Weeks since most recent delivery	8.2 (3.0)	9.9 (3.5)	7.1 (2.0)
Exclusively breastfeeding	51 (94.4%)	20 (95.2%)	31 (93.9%)
Nugent-BV (Nugent score ≥7)	26 (48.2%)	7 (33.3%)	19 (57.6%)
Vaginal washing[b]	30 (55.6%)	13 (61.9%)	17 (51.5%)
Any vaginal bleeding	12 (22.2%)	3 (14.3%)	9 (27.3%)
Use of antibiotics or metronidazole[c]	1 (1.9%)	0	1 (3.0%)
Resumed sexual intercourse since delivery	23 (42.6%)	6 (28.6%)	17 (51.5%)
Currently using condoms[d]	1 (4.4%)	1 (16.7%)	0

Abbreviations: BV, bacterial vaginosis; DMPA-IM, intramuscular depot-medroxyprogesterone acetate.
Enrollment: visit DMPA-IM was administered.

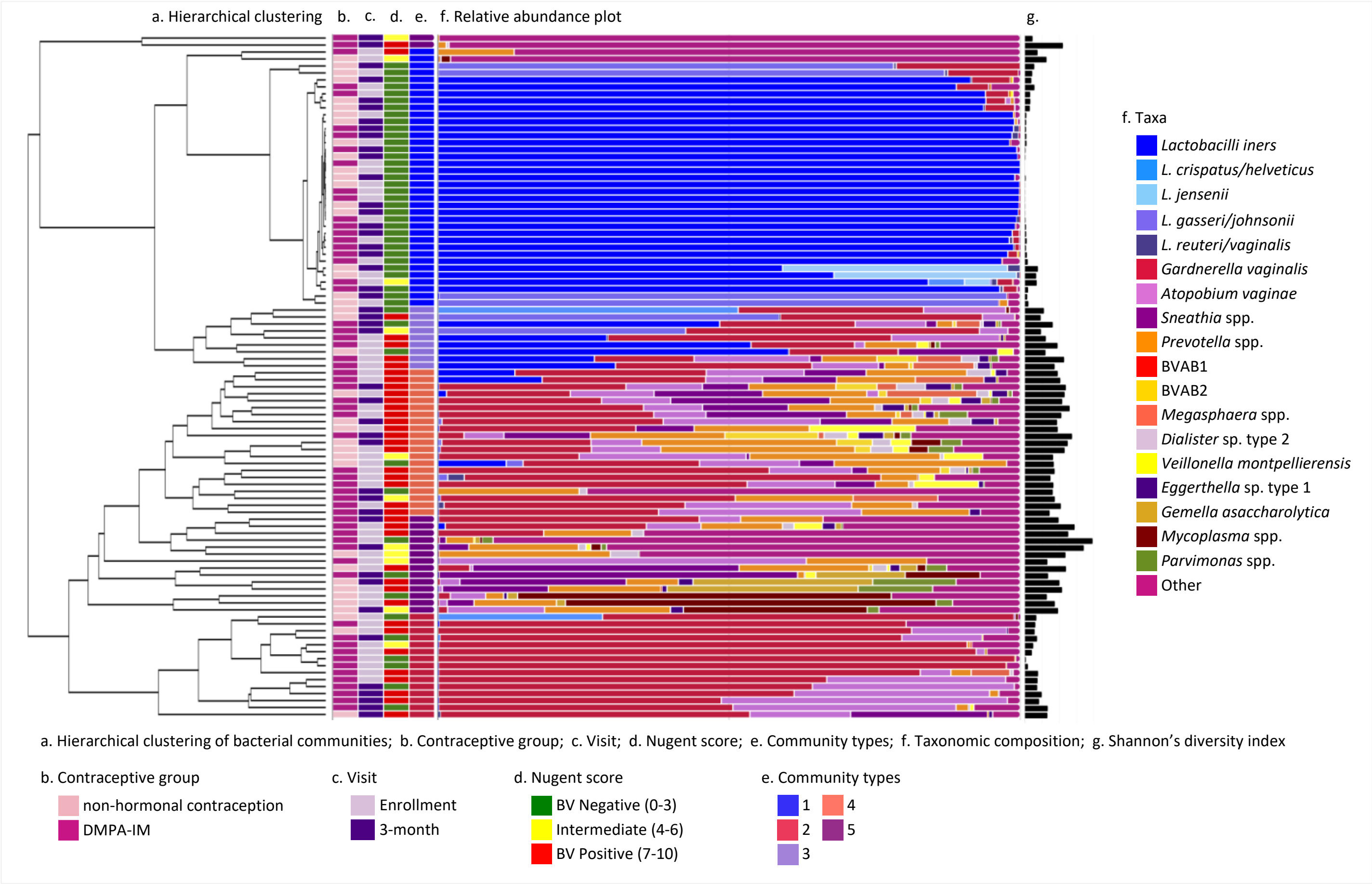
[a] Data presented as n (%) or mean (SD).

[b] Past week.

[c] Past month.

[d] Among those who had resumed sexual intercourse.

Figure 2.1. Vaginal bacterial communities among all women (N=98 swabs from 54 women). Unweighted pair group method with arithmetic mean (UPGMA) hierarchical clustering was used to generate a clustering tree depicting bacterial beta diversity. a. Scale bar represents UniFrac distance. Columns b-e represent: b. Contraceptive group, c. Study visit, d. Nugent-BV status, and e. Community types. f. Colored bars represent the relative abundance of bacterial taxa; less abundant taxa are grouped as "other". g. Black bars represent Shannon's diversity index.



Abbreviations: BV, bacterial vaginosis; DMPA-IM, intramuscular depot-medroxyprogesterone acetate; SDI, Shannon's diversity index. Enrollment: visit DMPA-IM was administered; Three-month: visit three-months post DMPA-IM administration.

Table 2.2. Description of the five vaginal bacterial community types identified among all women at enrollment and the three-month follow-up visit (N=98 swabs from 54 women).

Community type	**Total number of women**[a]	**Number of women with Nugent-BV in each CT**[b]	**Shannon's diversity index**[c]	**Taxa composition**[d]
CT1	37 (38%)	1 (3%)	0.24 (0.29)	Dominant taxa: *Lactobacillus iners* (79.7%)
CT2	15 (15%)	8 (53%)	0.69 (0.35)*	Dominant taxa: *Gardnerella vaginalis* (74.6%); prevalent taxa: *Atopobium vaginae* (16.1%)
CT3	9 (9%)	5 (56%)	1.44 (0.47)*	No dominant taxa; prevalent taxa: *G. vaginalis* (30.7%), *L. iners* (27.7%), *Lactobacillus gasseri/johnsonii* (11.2%)
CT4	21 (21%)	17 (81%)	2.00 (0.37)**	No dominant taxa; prevalent taxa: *G. vaginalis* (33.7%), *Prevotella* species (18.2%), *A. vaginae* (13.2%)
CT5	16 (16%)	8 (50%)	2.12 (0.85)	No dominant taxa: Heterogenous mix of anaerobic bacteria
All	98 (100%)	39 (40%)	1.10 (0.93)	Not applicable

Abbreviations: BV, bacterial vaginosis; CT, community type.

[a] Data presented as n (%); percents may not sum to 100 due to rounding.

[b] Data presented as n (%) of women in cluster.

[c] Data presented as mean (SD).

[d] Percent displayed is mean relative abundance for that taxa in members of corresponding community type.

Significantly more diverse (as measured by Two sample t-test comparing Shannon's diversity index) than the preceding CT at the *$p<0.001$ level and **$p<0.01$ level.

Table 2.3. Comparison of Shannon's diversity index and the five community types by contraceptive group at enrollment and the three-month follow-up visit.

	Enrollment Visit			**Three-month Follow-up Visit**		
Characteristic[a]	**Non-hormonal contraception (n=21)**	**DMPA-IM (n=33)**	**P-value**[b]	**Non-hormonal contraception (n=18)**	**DMPA-IM (n=26)**	**P-value**[b]
Community type			0.54			0.27
CT 1	8 (38.1%)	10 (30.3%)		11 (61.1%)	8 (30.8%)	
CT 2	2 (9.5%)	7 (21.2%)		1 (5.6%)	5 (19.2%)	
CT 3	1 (4.8%)	4 (12.1%)		2 (11.1%)	2 (7.7%)	
CT 4	5 (23.8%)	8 (24.2%)		2 (11.1%)	6 (23.1%)	
CT 5	5 (23.8%)	4 (12.1%)		2 (11.1%)	5 (19.2%)	
Community type			0.55			**0.046**
Lactobacilli dominant (CT1)	8 (38.1%)	10 (30.3%)		11 (61.1%)	8 (30.8%)	
Lactobacilli deplete (CTs 2-5)	13 (61.9%)	23 (69.7%)		7 (38.9%)	18 (69.2%)	
Nugent score	3.62±3.75	5.30±3.75	0.11	2.67±3.68	3.35±3.70	0.55
Shannon's diversity index	1.19±0.87	1.19±1.01	0.99	0.79±0.81	1.13±0.96	0.23
Total bacterial load[c]	8.31±0.84	8.59±0.58	0.16	8.74±0.50	8.31±0.89	**0.07**

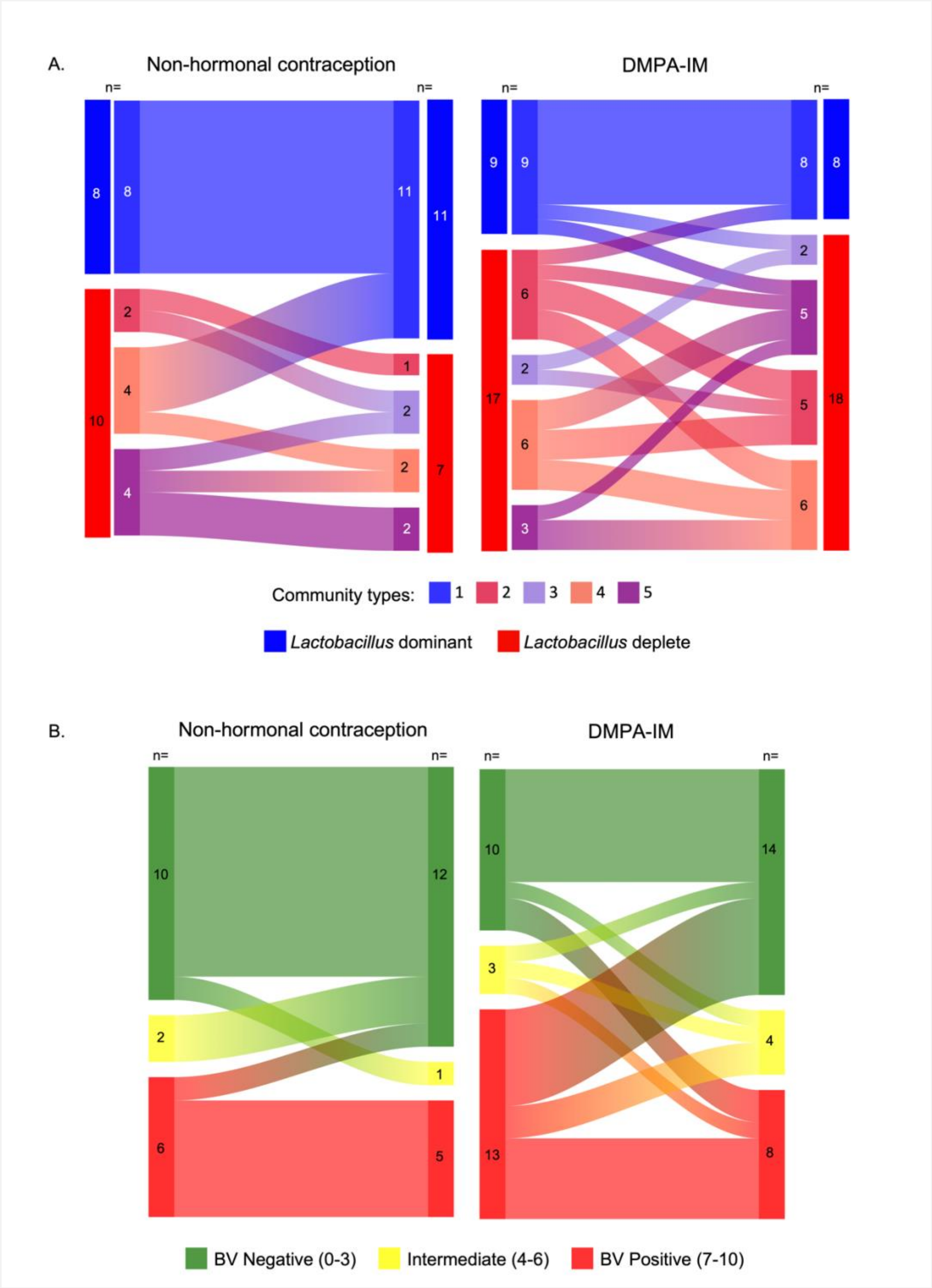
Abbreviations: CT, community type; DMPA-IM, intramuscular depot-medroxyprogesterone acetate.
Enrollment: visit DMPA-IM was administered; Three-month: visit three-months post DMPA-IM administration.
Bolded text indicates significant values.

[a] Data presented as n (%) or mean (SD).

[b] Chi-square test for categorical variables and unpaired t-test for continuous variables.

[c] Displayed as $\log_{10}$ 16S rRNA gene copies per swab.

Figure 2.2. Change in (A) community type membership and (B) Nugent-BV category from enrollment to the three-month follow-up visit for non-hormonal contraception users (n=18) and DMPA-IM users (n=26) among the 44 women who returned for follow-up.



Abbreviations: BV, bacterial vaginosis; CT, community type; DMPA-IM, intramuscular depot-medroxyprogesterone acetate. Enrollment: visit DMPA-IM was administered; Three-month: visit three-months post DMPA-IM administration.

Table 2.4. Change in Nugent score, total bacterial load, and Shannon's diversity index from enrollment to the three-month follow-up visit by contraceptive group, and difference in change between contraceptive groups.

	Non-hormonal contraception	95% CI	P-value	DMPA-IM	95% CI	P-value	Difference	95% CI	P-value
Nugent score	-0.73	-2.18, 0.73	0.327	-1.89	-3.53, -0.25	**0.024**	-1.17	-3.35, 1.02	0.297
Shannon's diversity index	-0.34	-0.67, -0.001	**0.050**	0.03	-0.24, 0.30	0.831	0.37	-0.07, 0.80	0.098
Total bacterial load[a]	0.37	-0.02, 0.76	0.062	-0.26	-0.65, 0.13	0.189	-0.64	-1.19, -0.08	**0.024**

Abbreviations: CI, confidence interval; DMPA-IM, intramuscular depot-medroxyprogesterone acetate.

Data presented as mean change (95% CI).

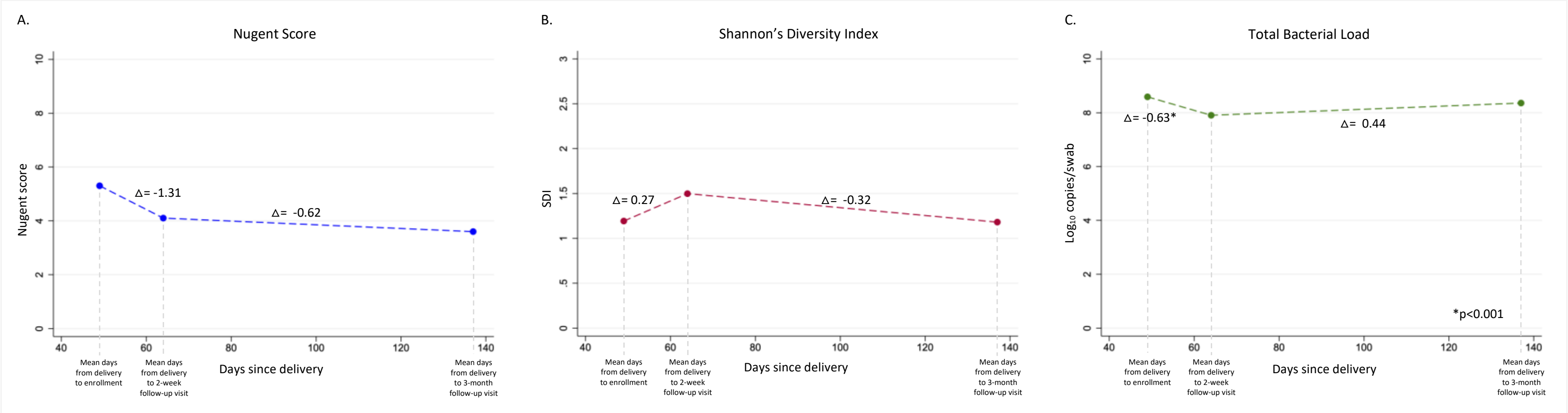
Enrollment: visit DMPA-IM was administered; Three-month: visit three-months post DMPA-IM administration.

Results from linear mixed models adjusted for time (days from delivery to enrollment and days from enrollment/DMPA-IM injection to swab collection), age, marital status, and resumption of intercourse ate enrollment.

Bolded text indicates significant values.

[a] Displayed as $\log_{10}$ 16S rRNA gene copies per swab.

Figure 2.3. Mean change in (A) Nugent score, (B) Shannon's diversity index, and (C) total bacterial load from enrollment to two-weeks post-injection and from two-weeks post-injection to the three-month follow-up visit among DMPA-IM users only.



	(A) Nugent score	**(B) Shannon's diversity index**	**(C) Total bacterial load**[a]
Enrollment to two-week visit	=-1.31 (95%CI:-2.77, 0.15); p=0.078	=0.27 (95%CI: -0.11, 0.64); p=0.163	=-0.63 (95%CI: -0.98, -0.28); **p<0.001**
Two-week visit to three-month visit	=-0.62 (95%CI: -2.15, 0.91); p=0.426	=-0.32 (95%CI: -0.76, 0.12); p=0.155	=0.44 (95%CI: -0.05, 0.95); p=0.081
Difference	=0.69 (95%CI: -1.87, 3.24); p=0.598	=-0.58 (95%CI: -1.35, 0.19); p=0.137	=-1.07 (95%CI: 0.29, 1.85); **p=0.007**

Abbreviations: CI, confidence interval; DMPA-IM, intramuscular depot-medroxyprogesterone acetate.

Data presented as mean change (95% CI).

Enrollment: visit DMPA-IM was administered; two-week: visit two-weeks post DMPA-IM administration (DMPA users only); Three-month: visit three-months post DMPA-IM administration.

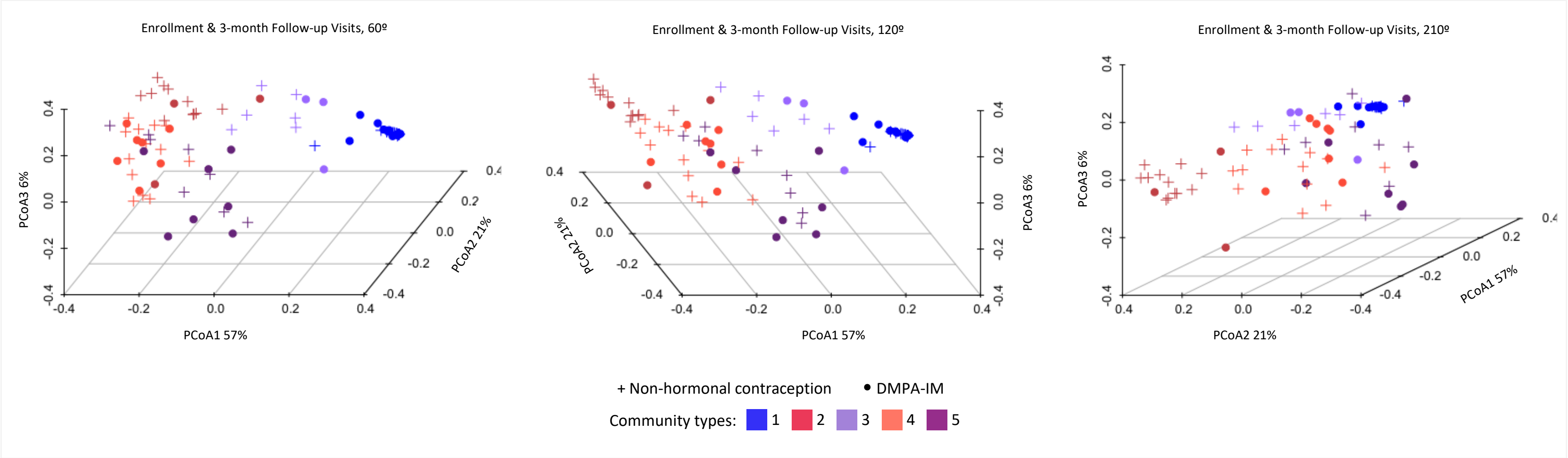
Results from linear mixed models adjusted for time (days from delivery to enrollment and days from enrollment/DMPA-IM injection to swab collection).

Bolded text indicates significant values.

[a] Displayed as $\log_{10}$ 16S rRNA gene copies per swab.

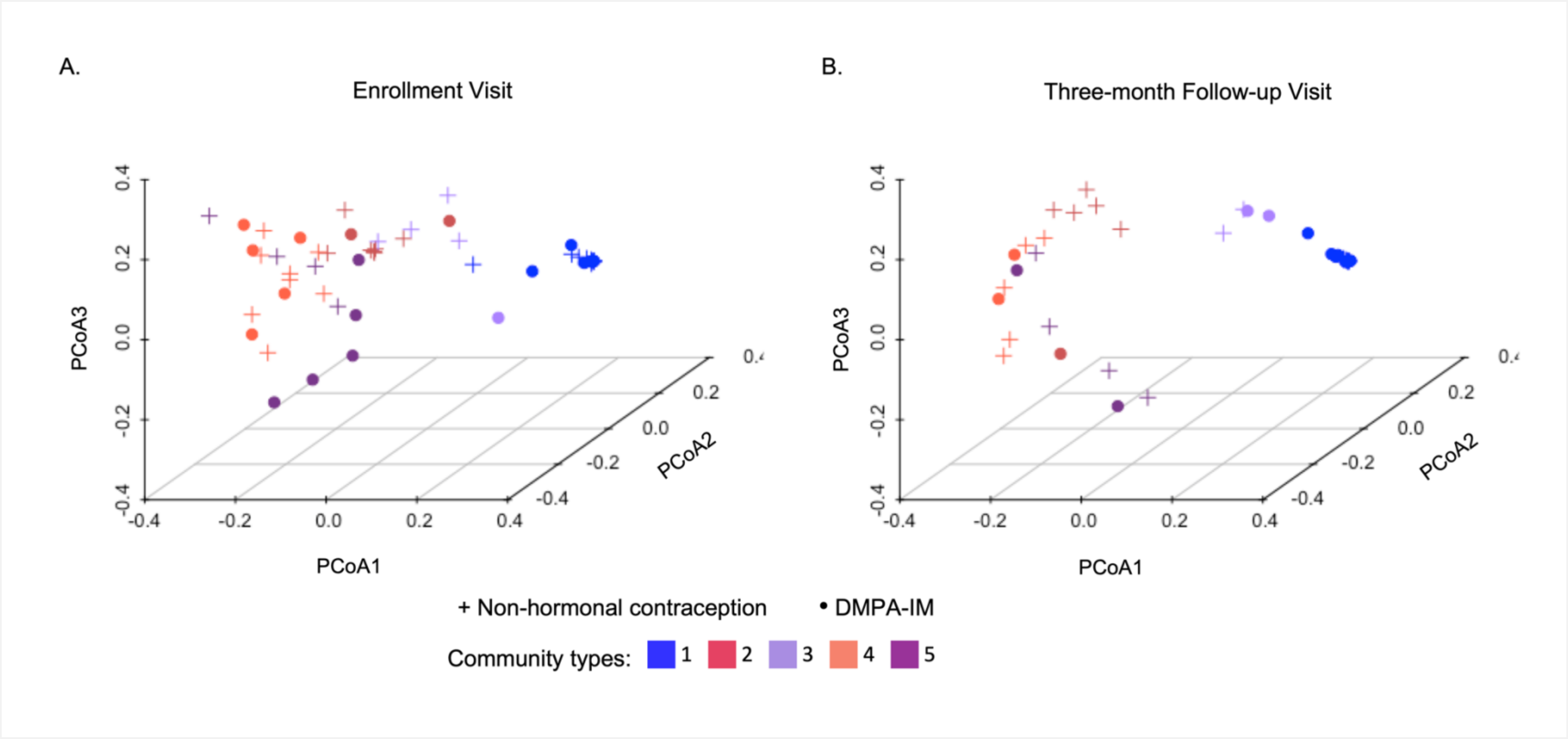
SUPPLEMENTAL TABLES AND FIGURES

Figure S2.1. Principal coordinates analysis plots of the 98 observations, distinguished by contraceptive group and community types.



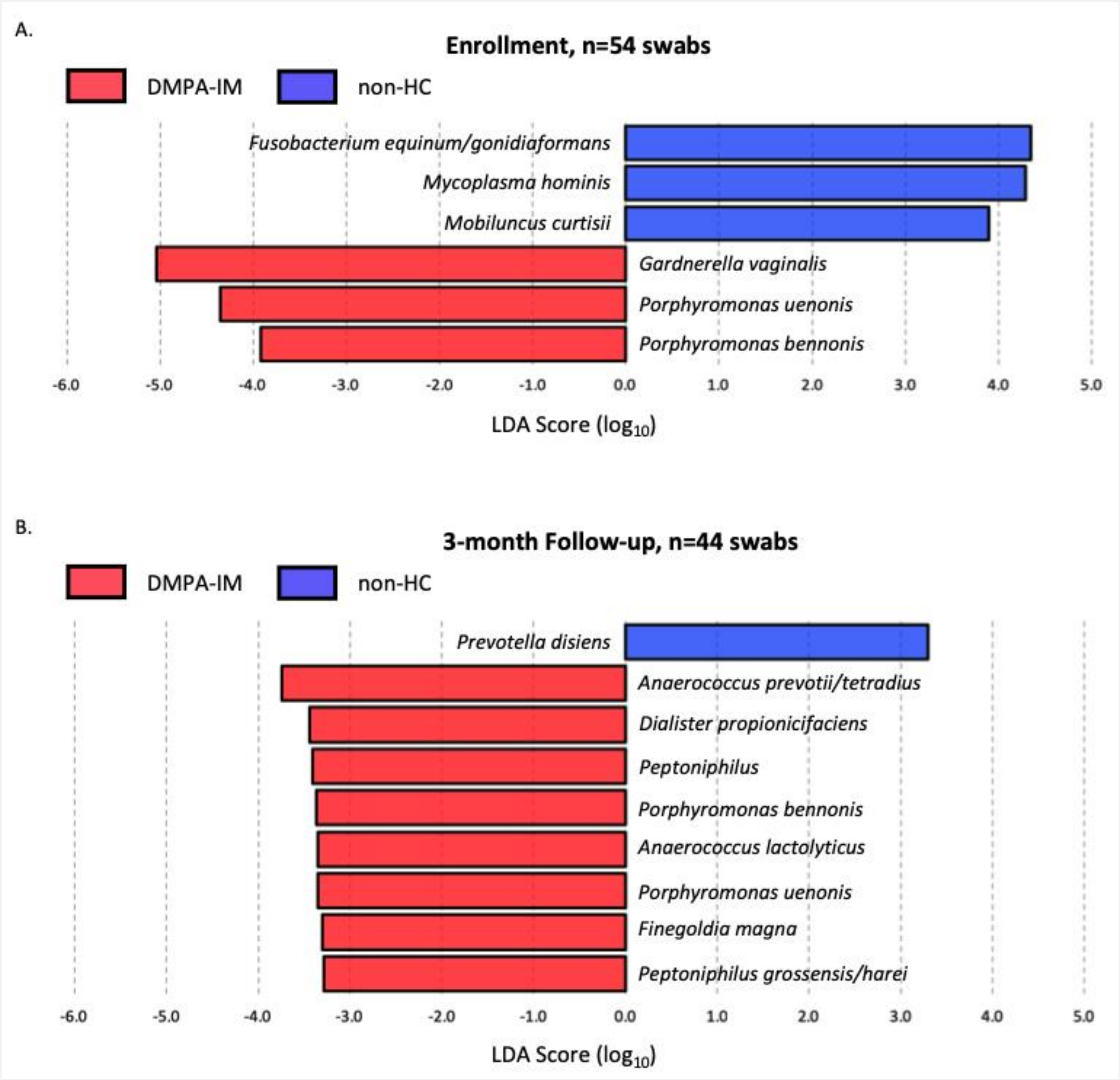
Abbreviations: DMPA-IM, intramuscular depot-medroxyprogesterone acetate; PCoA, principal coordinate.
Enrollment: visit DMPA-IM was administered; Three-month: visit three-months post DMPA-IM administration.

Figure S2.2. Principal coordinates analysis plots, distinguished by contraceptive group and community types, at: (A) enrollment (n=54) and (B) the three-month follow-up visit (n=44).



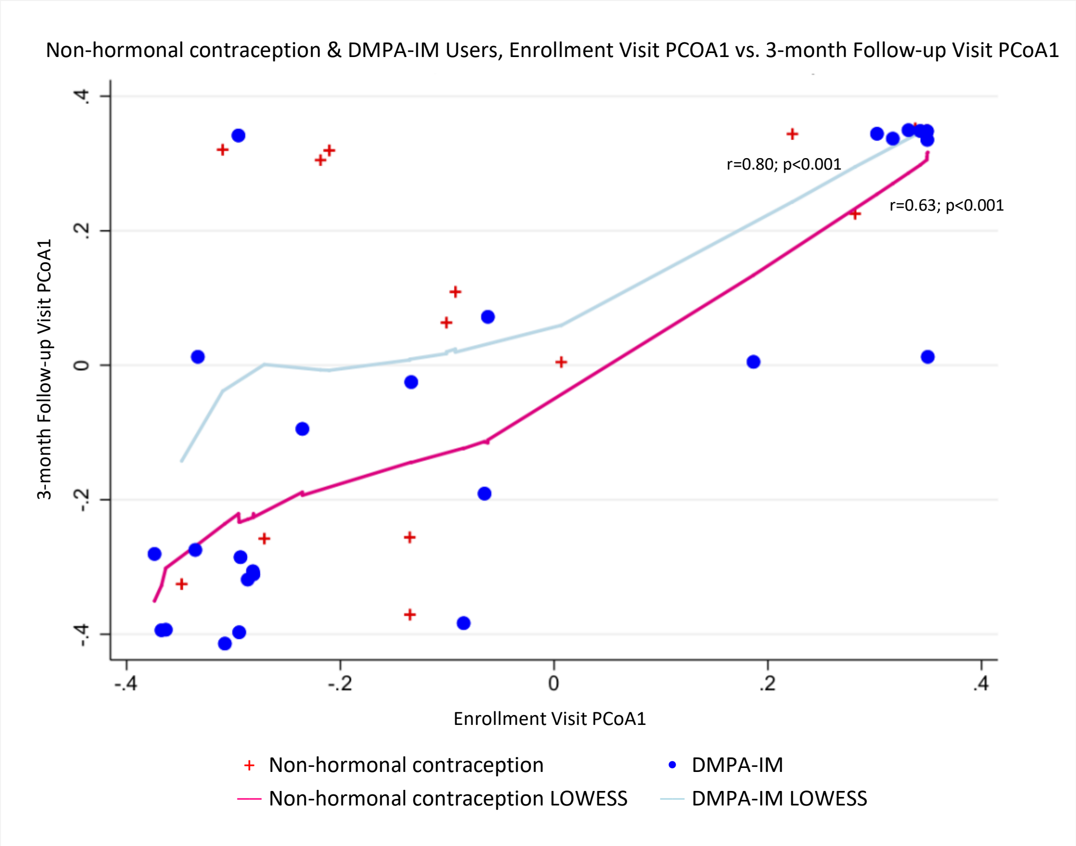
Abbreviations: DMPA-IM, intramuscular depot-medroxyprogesterone acetate; PCoA, principal coordinate.
Enrollment: visit DMPA-IM was administered; Three-month: visit three-months post DMPA-IM administration.

Figure S2.3. Linear discriminant analysis with effect size (LEfSe) results indicating significant differences (α=0.05) in the relative abundance of specific bacterial taxa between contraceptive groups; taxa that were more abundant among DMPA-IM users are shown in red and taxa that were more abundant among non-hormonal contraception users are shown in blue at: (A) enrollment among all women (n=54) and (B) the three-month follow-up visit (n=44).



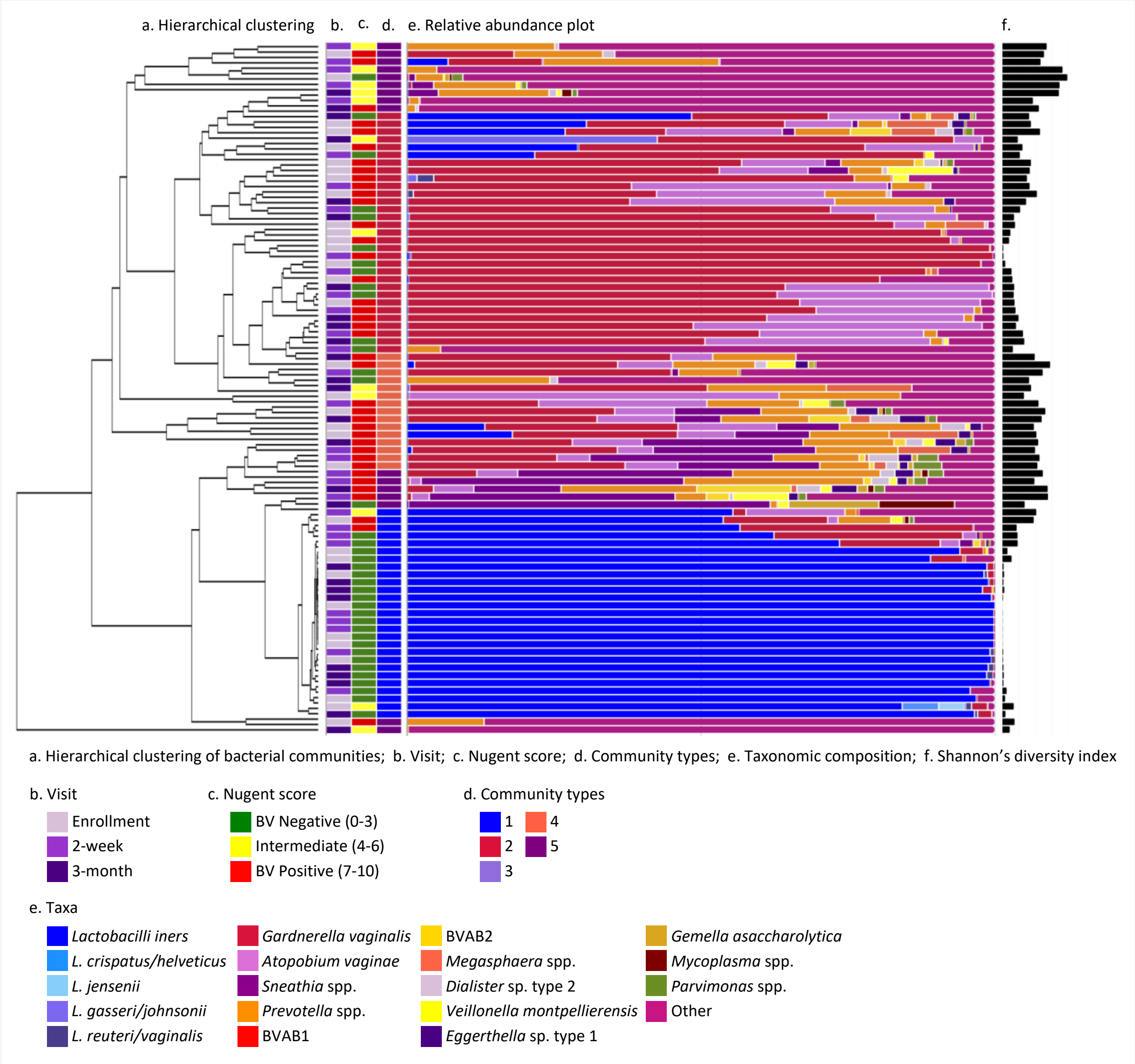
Abbreviations: DMPA-IM, intramuscular depot-medroxyprogesterone acetate; HC, hormonal contraception.
Enrollment: visit DMPA-IM was administered; Three-month: visit three-months post DMPA-IM administration.

Figure S2.4. Plot of the enrollment visit's first principal coordinate vs the three-month follow-up visit's first principal coordinate, distinguished by contraceptive group.



Abbreviations: DMPA-IM, intramuscular depot-medroxyprogesterone acetate; LOWESS: Locally Weighted Scatterplot Smoothing; PCoA, principal coordinate.
Enrollment: visit DMPA-IM was administered; Three-month: visit three-months post DMPA-IM administration.

Figure S2.5. Vaginal bacterial communities among women using DMPA-IM (N=88 swabs from 33 women). Unweighted pair group method with arithmetic mean (UPGMA) hierarchical clustering was used to generate a clustering tree depicting bacterial beta diversity. a. Scale bar represents UniFrac distance. Columns b-d represent: b. Study visit, c. Nugent-BV status, and d. Community types. e. Colored bars represent the relative abundance of bacterial taxa; less abundant taxa are grouped as “other”. f. Black bars represent Shannon’s diversity index.



Abbreviations: BV, bacterial vaginosis; DMPA-IM, intramuscular depot-medroxyprogesterone acetate; SDI, Shannon’s diversity index.
Enrollment: visit DMPA-IM was administered; two-week: visit two-weeks post DMPA-IM administration (DMPA users only); Three-month: visit three-months post DMPA-IM administration.

Table S2.1. Description of the four vaginal bacterial community types identified among DMPA-IM users at all visits (N=89 swabs from 33 women).

Community type	Total number of women[a]	Number of women with Nugent-BV in each CT[b]	Shannon's diversity index[c]	Taxa composition[d]
CT1	27 (30.3%)	2 (7.4%)	0.36 (0.55)	Dominant taxa: *Lactobacillus iners* (90.6%)
CT2	31 (34.8%)	18 (58.1%)	1.03 (0.57)*	Dominant taxa: *Gardnerella vaginalis* (61.4%); prevalent taxa: *Atopobium vaginae* (16.8%)
CT4	15 (16.9%)	11 (73.3%)	2.11 (0.39)*	No dominant taxa; prevalent taxa: *G. vaginalis* (30.1%), *Prevotella* species (15.5%), *A. vaginae* (12.7%), *Sneathia* species (10.5%)
CT5	16 (18.0%)	8 (50.0%)	2.40 (0.99)	No dominant taxa: Heterogenous mix of anaerobic bacteria
All	89 (100%)	39 (43.8%)	1.25 (1.01)	Not applicable

Abbreviations: BV, bacterial vaginosis; CT, community type; DMPA-IM, intramuscular depot-medroxyprogesterone acetate.

[a] Data presented as n (%); percents may not sum to 100 due to rounding.

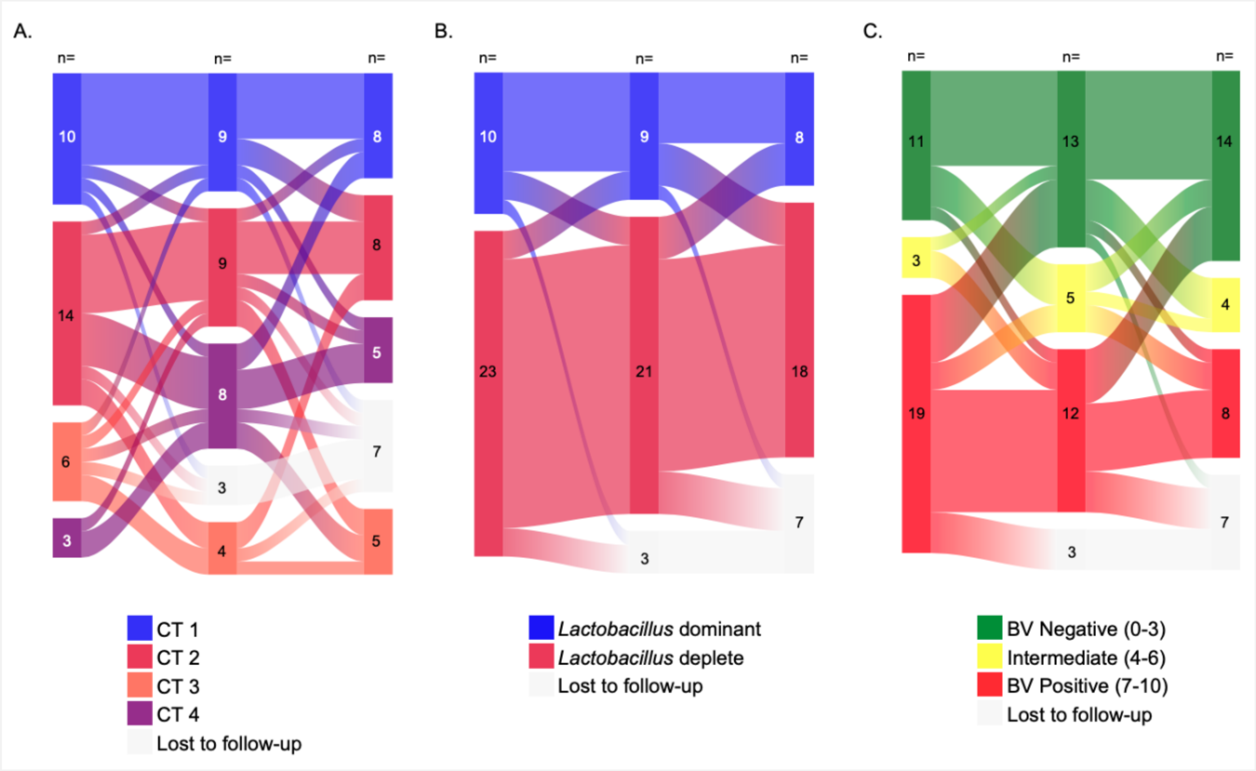
[b] Data presented as n (%) of women in cluster.

[c] Data presented as mean (SD).

[d] Percent displayed is mean relative abundance for that taxa in members of corresponding community type.

Significantly more diverse (as measured by Two sample t-test comparing Shannon's diversity index) than the preceding CT at the *p<0.001 level.

Figure S2.6. Change in (A) community type membership, (B) *Lactobacillus* dominance, and (C) Nugent-BV category, from enrollment to two-weeks post-injection and from two-weeks post-injection to the three-month follow-up visit among DMPA-IM users.



Abbreviations: BV, bacterial vaginosis; CT, community type; DMPA-IM, intramuscular depot-medroxyprogesterone acetate. Enrollment: visit DMPA-IM was administered; two-week: visit two-weeks post DMPA-IM administration (DMPA users only); Three-month: visit three-months post DMPA-IM administration.

Table S2.2. Analysis of dose-response relationship between plasma MPA level categorized into tertiles (measured at the two-week follow-up visit) and Nugent score, total bacterial load, and Shannon's diversity index among DMPA-IM users only at: (1) the two-week follow-up visit (n=28) and (2) the 3-month follow-up visit (n=26).

Two-week follow-up visit (n=28)									
Plasma MPA Categories (ng/ml)	**Nugent score**	**95% CI**	**P-value**	**Shannon's diversity index**	**95% CI**	**P-value**	**Total bacterial load**[a]	**95% CI**	**P-value**
0.69-1.09 (Ref)	0	--	--	0	--	--	0	--	--
1.10-1.50	-1.25	-4.69, 2.20	0.463	-0.39	-1.30, 0.51	0.381	0.09	-1.18, 1.35	0.886
1.51-3.54	-0.32	-3.79, 3.14	0.848	-0.30	-1.40, 0.81	0.585	-0.42	-1.67, 0.82	0.489

Three-month follow-up visit (n=26)									
Plasma MPA Categories (ng/ml)	**Nugent score**	**95% CI**	**P-value**	**Shannon's diversity index**	**95% CI**	**P-value**	**Total bacterial load**[a]	**95% CI**	**P-value**
0.69-1.09 (Ref)	0	--	--	0	--	--	0	--	--
1.10-1.50	0.92	-2.88, 4.72	0.621	-0.15	-1.04, 0.74	0.729	-0.19	-1.04, 0.65	0.640
1.51-3.54	0.34	-4.22, 4.91	0.878	0.23	-0.91, 1.37	0.684	-0.19	-0.83, 0.46	0.553

Abbreviations: CI, confidence interval; DMPA-IM, intramuscular depot-medroxyprogesterone acetate; MPA, medroxyprogesterone acetate.

Enrollment: visit DMPA-IM was administered; two-week: visit two-weeks post DMPA-IM administration (DMPA users only); Three-month: visit three-months post DMPA-IM administration.

Results from linear mixed models adjusted for time (days from delivery to enrollment and days from enrollment/DMPA-IM injection to swab collection).

[a] Displayed as $\log_{10}$ 16S rRNA gene copies per swab.

CHAPTER 3. Changes in key vaginal bacteria among postpartum African women initiating intramuscular depot-medroxyprogesterone acetate

ABSTRACT

Background: The ECHO trial has relieved apprehension about intramuscular depot medroxyprogesterone acetate (DMPA-IM), however it is still important to understand how DMPA-IM affects the vaginal environment. We sought to describe how DMPA-IM initiation influences vaginal bacteria associated with HIV acquisition in postpartum women.

Methods: Vaginal swabs were collected for Nugent score determination and taxon-specific quantitative PCR of eight bacteria. Enrollment occurred at contraceptive initiation (DMPA-IM or non-hormonal contraception (non-HC)) and repeat vaginal swabs were collected after three months. Generalized estimating equations were used to estimate changes in Nugent score, total bacterial load, and taxa concentrations among contraceptive groups.

Results: Women who chose DMPA-IM (n=33) were more likely to be married (97%vs.67%) and have resumed intercourse since delivery (52%vs.29%) compared to women who chose non-HC (n=21). After three months, significant decreases in the concentrations of *Sneathia* species, *Mycoplasma hominis*, and *Parvimonas* species Type 1 were seen among non-HC users, however concentrations remained stable among DMPA-IM users; contraceptive method was associated with significantly different changes in *M. hominis* concentration between groups (p=0.010).

Conclusion: Our findings suggest that postpartum use of DMPA-IM and non-HC may have differential impacts on the vaginal concentrations of some bacteria that have previously been associated with HIV acquisition.

INTRODUCTION

The vaginal microbiome plays a key role in women's reproductive health. Hydrogen peroxide-producing *Lactobacillus*-dominant vaginal bacterial communities are considered optimal for health [11,12]. Communities dominated by anaerobic bacteria, or non-optimal microbiota, as well as specific microbial taxa, are associated with subclinical inflammation, poor reproductive health outcomes, and sexually transmitted infections [14-17]. Hormones, including estrogens and progestogens, play important roles in vaginal microbial ecology [22-24], and exogenous hormones, such as contraceptives, may induce important changes in the composition of vaginal microbiota and production of soluble factors by bacteria.

Injectable progestin-only contraceptives, including intramuscular depot-medroxyprogesterone acetate (DMPA-IM), are the most commonly used contraceptives among women in sub-Saharan Africa [69], where the burden of vaginal dysbiosis is highest [3] and 25.7 million people live with HIV [4]. DMPA-IM inhibits the secretion of pituitary gonadotropins, resulting in anovulation and decreased production of estrogen [6]. Reduced levels of estrogen have been associated with glycogen suppression [23,25], and glycogen is an important substrate for *Lactobacillus* species [25-27]. Consequently, a reduction of glycogen in the female reproductive tract could shift the vaginal microbiome to an anaerobic, non-optimal state. Changes in vaginal bleeding patterns and perturbation of the regular menstrual cycle associated with DMPA-IM initiation may also impact the vaginal microbiome [24,28]. Additionally, previous observational research has suggested that DMPA-IM use may be associated with increased risk of HIV acquisition [29,70-72]. While recently released findings from The Evidence for Contraceptive Options and HIV Outcomes (ECHO) Trial have generally reduced clinical concern about DMPA-

IM use in women at risk of HIV [31], there is still uncertainty in the field as to whether DMPA-IM increases the risk of HIV acquisition relative to non-HC methods [32]. Therefore, it is important to understand how DMPA-IM affects the vaginal environment, including vaginal microbiota.

Several anaerobic bacteria have been identified as significantly associated with increased HIV acquisition, including: *Prevotella* species (*P. melaninogenica* [16], *P. bivia* [16]), *Mycoplasma* species (*Mycoplasma* spp. [16], *M. hominis* [17]), *Sneathia* species (*S. sanguinegens* [16], *Leptotrichia/Sneathia* (now *Sneathia* spp.) [17]), *Parvimonas* species type 2 [17], *Gemella asaccharolytica* [17], *Eggerthella* species Type 1 [17], and *Megasphaera* species (*Megasphaera* spp. Types 1 and 2 (combined assay) [17]). Observed heterogeneity in species-level determination of *Mycoplasma* and *Sneathia* species between studies may be due, at least in part, to different laboratory and bioinformatics methods [46]. In addition, unpublished data from two studies have also linked *P. bivia* with increased HIV acquisition risk [73,74], and one of the studies reported associations between *M. hominis*, *Leptotrichia/Sneathia* (now *Sneathia* spp.), *Parvimonas* sp. type 2, *G. asaccharolytica*, *Eggerthella* sp. Type 1, and *Megasphaera* sp. Type 1 and increased HIV acquisition risk [74], replicating the published findings from other East African cohorts [17].

A limited number of studies have evaluated DMPA-IM's effect on the overall vaginal microbiome [37,38,40-43], and to our knowledge there are no published studies specifically evaluating the effect of DMPA-IM on the concentration of a majority of bacterial taxa associated with HIV acquisition. Additionally, there is a gap in this data for postpartum women. Therefore, we sought to investigate if DMPA-IM use is associated with increases in bacteria previously associated with HIV acquisition [16,17,73,74]. Based on pilot data from a cohort of Kenyan women [40], we hypothesized that quantities of *G. vaginalis* would decline in DMPA-IM users and other anaerobic bacteria would increase to fill the opening.

METHODS

Study setting, subjects, and design

We designed a prospective cohort study of postpartum women to assess how DMPA-IM affects vaginal environment, including alterations to vaginal microbiota, in the three months following DMPA-IM initiation. Breastfeeding women, 6-14 weeks postpartum who sought contraception counseling at a public primary care clinic in Nairobi, Kenya were recruited for enrollment. Because the menstrual cycle is associated with changes in vaginal microbiota [24,28,67], enrollment was limited to lactating, amenorrheic women to reduce expected variability at baseline. Women were eligible for enrollment if they were HIV-negative and chose DMPA-IM or non-HC (condoms, lactational amenorrhea, rhythm) as their contraceptive method. Women already using a hormonal contraceptive or unwilling to learn their HIV status were ineligible. Women with evidence of cervicitis or STI at enrollment were excluded.

Women attended two or three study visits depending on contraceptive method. At the enrollment visit women chose their contraceptive method; those who chose DMPA-IM received an injection from study staff, verifying the exact time of DMPA-IM administration. DMPA-IM users were asked to return approximately 9-14 days later for a brief visit coinciding with typical peak plasma medroxyprogesterone acetate (MPA) concentrations, the active component of DMPA-IM [51,75]. All women were asked to return for a follow-up visit 3 months post-enrollment. Demographic, health, and sexual activity information were collected via questionnaires at each visit. Vaginal swabs for microbiota evaluation were collected during pelvic exams; swabs were taken bilaterally in the space that includes right and left lateral fornix and the distal two-thirds of the vaginal wall. At the enrollment visit, all specimens were collected before DMPA-IM administration. Swabs were not collected from women with vaginal bleeding, including any spotting related to DMPA-IM initiation. If women had any vaginal bleeding or spotting, they were asked to return for swab collection when bleeding had stopped.

Vaginal swabs were cryopreserved at –80°C and batch shipped on dry ice to the Fred Hutchinson Cancer Research Center (Seattle, WA, USA) for analysis.

Laboratory procedures

Peripheral blood was used for point-of-care HIV testing (Determine rapid enzyme-linked immunosorbent assay, Abbott, Abbott Park, Illinois). Vaginal saline wet mounts were examined microscopically for the presence of motile trichomonads and fungal elements. Gram stains of vaginal fluid were used for evaluation of bacterial vaginosis (BV), with Nugent scores of ≥7 considered a diagnosis of Nugent-BV (method of Nugent and Hillier [55]).

Quantitative PCR (qPCR) was performed for eight bacterial taxa (*Gardnerella vaginalis*, *M. hominis*, *Sneathia* species, *G. asaccharolytica*, *Eggerthella* sp. Type 1, *Megasphaera* spp. Types 1 and 2 (combined assay), *Parvimonas* sp. Type 1, and *Parvimonas* sp. Type 2), using previously described methods [17,76,77]. Detailed methodology on DNA extraction and amplification have been detailed previously [17,56,57,78] and are described in the supplemental material.

Statistical analyses

Bacterial DNA concentrations were $\log_{10}$-transformed to normalize their distribution. DNA concentrations, when not detected, were assigned a value of half the lower limit of detection (LLD) of the assay (1.495 $\log_{10}$ copies/swab for all taxa except *Eggerthella* sp. Type 1 which had a LLD of 1.796 $\log_{10}$ copies/swab).

Our primary analysis was evaluation of changes in Nugent score, total bacterial load, and the concentration of the eight selected bacterial taxa among DMPA-IM users and non-HC users between enrollment and follow-up, as well as to compare changes over time between the contraceptive groups. Change from enrollment to the three-month follow-up visit among all women was estimated separately for each outcome using generalized estimating equations (GEE) with an interaction term between contraceptive group and time (days from enrollment/DMPA-IM injection to swab collection). Models including all women were adjusted for time from delivery to enrollment, as well as important potential confounders selected *a priori*, including age, marital status, and resumption of intercourse at enrollment. Patterns of change over time were visualized using Spaghetti plots, with trend lines for the mean change in the outcomes for each contraceptive group estimated from the GEE models.

Due to a large proportion of women having undetectable concentrations of the eight selected bacterial taxa, we also examined the pattern of taxon detection over follow-up and evaluated change in detectability of each taxa by contraceptive group using McNemar's test. Sensitivity analyses were performed on the subsets of participants with at least one detectable value for that taxon during follow-up (women with a value above the lower limit of detection for a specific taxon at one or both visits). In these subsets, we only adjusted for time from delivery to enrollment since a majority of the taxa were detectable in 20 women or less and we did not want to overfit the models [79]. Patterns of change were visualized using Spaghetti plots, with trend lines for mean change among women with at least one detectable value for that taxon estimated from the GEE models.

Within the DMPA-IM group, we assessed whether the time of peak serum MPA concentrations was associated with changes in the outcomes using similar methods as above. We evaluated patterns of change over follow-up, including change from enrollment to the two-week post-injection visit, when MPA serum concentrations are approximately at peak by published pharmacokinetic data [51,75], and from the two-week post-injection visit to the three-month follow-up visit.

For all associations, the significance level was set at $p<0.05$. Analyses were performed using Stata version 14 (StataCorp, College Station TX). The research protocol was approved by the Kenyatta National Hospital Ethics and Research Committee and the University of Washington Institutional Review Board. Written informed consent was obtained in English or Kiswahili from all participants.

RESULTS

Participant characteristics

We enrolled 54 women, 33 (61%) of whom chose to initiate DMPA-IM. Women who chose DMPA-IM sought contraception counseling sooner than women who chose non-HC methods (7.1 weeks (standard deviation (SD): 2.0) post-delivery vs. 9.9 weeks (SD: 3.5) post-delivery) (Table 3.1). A higher proportion of DMPA-IM users were married (97% vs. 67%) and had resumed sexual intercourse by enrollment (52% vs. 29%) (Table 3.2). There was a non-significant trend towards increased Nugent-BV at baseline among women who chose DMPA-IM compared to those who chose non-HC (unadjusted Odds Ratio (OR)=2.71, 95% Confidence Interval (CI): 0.83-8.87); this trend was not present at follow-up. Vaginal washing was common (DMPA-IM: 52%, non-HC: 62%) and became more prevalent in both groups, with a larger increase among DMPA-IM users (DMPA-IM: 85%, non-HC: 72%). Of the enrolled women, 44 (81%) women returned for their three-month follow-up visit. Time between the enrollment visit and the three-month follow-up visit was comparable between the two contraceptive groups (84 days for non-HC users [interquartile range (IRQ): 84-90 days] and 84 days for DMPA-IM users [IRQ: 84-89 days]). There were no incident HIV infections during follow-up.

Detection of the 8 taxa in participants

All eight taxa assessed were found in this group of postpartum women, however only *G. vaginalis* was detectable in a majority of participants at enrollment (non-HC: 71%; DMPA-IM: 85%) and follow-up (non-HC: 85%; DMPA-IM: 73%); the other seven taxa were detectable in fewer than 50% of women (Table 3.3). There was no difference in the detectability of the eight taxa between the two contraceptive groups at enrollment, nor was there any significant difference in mean concentration of each taxon among women with the taxon detected, with the exception of *G. asaccharolytica* which was found at a slightly higher concentration among non-HC users at enrollment (p=0.037) (Table 3.3). Women had an average of 3.0 taxa (SD 2.5) present at enrollment and 2.5 (SD 2.4) at the three-month follow-up visit, neither of which were significantly different between contraceptive groups (p=0.609 and p=0.653, respectively) or from baseline to follow-up for either contraceptive group (non-HC: p=0.253; DMPA-IM: p=

0.678) (Figure S3.1). Within contraceptive group, detection of each taxon at follow-up was not significantly different from detection at enrollment (Table S3.1). Qualitatively, very few women went from undetectable to detectable, or vice-versa, from enrollment to follow-up for each of the eight selected bacterial taxa (Figure 3.1).

Patterns of vaginal microbiota change among intramuscular depot-medroxyprogesterone acetate users compared to non-hormonal contraceptive users

Over three months, mean Nugent score decreased by 1.90 points (95%CI: -3.38 to -0.41, p=0.012) among women using DMPA-IM and by 0.73 points (95%CI: -2.70 to 1.24, p=0.469) among women using non-HC, resulting in similar mean Nugent scores in both contraceptive groups at follow-up (Table 3.4, Figure 3.2A). The difference in change between the two contraceptive groups was not statistically significant, however (p=0.354). In a small subset of women, the opposite pattern was observed: these women maintained an elevated Nugent score or increased in Nugent score after enrollment. Maintenance of an elevated Nugent score or an increase in Nugent score after enrollment was observed mostly among women using DMPA-IM (n=5 [19%] DMPA-IM users vs. n=2 [11%] non-HC users). Non-HC users started with lower Nugent scores, and were less likely to change Nugent category (Figure S3.2). Total bacterial load decreased non-significantly among DMPA-IM users and increased non-significantly among non-HC users over follow-up, resulting in a significant difference in changes of total bacterial load between the two contraceptive groups (difference in $\triangle$=-0.58 $\log_{10}$ copies/swab [95%CI: 1.16 to -0.01], p=0.049) (Table 3.4, Figure 3.2B).

In analyses including all women, the mean concentration of *G. vaginalis* decreased by 0.87 $\log_{10}$ copies/swab (95%CI: -1.81 to 0.06, p=0.067) among DMPA-IM users and increased by

0.22 (95%CI: -1.02 to 1.46, p=0.725) among non-HC users (Table 3.4, Figure 3.2C). As was seen with Nugent score, there was a small subset of DMPA-IM users (n=5 [19%]) who did not follow the overall trend; this subset was characterized by an increase in *G. vaginalis* concentration of ≥1 $\log_{10}$ copies/swab. Of these 5 women with a pattern of increasing *G. vaginalis,* 3 (60%) also exhibited the pattern of increasing Nugent score.

The concentration of three of the bacterial taxa assessed decreased significantly among women using non-HC over follow-up. *Sneathia* spp. (△=-1.09 $\log_{10}$ copies/swab [95%CI: -2.05 to -0.13], p=0.026), *M. hominis* (△=-1.21 $\log_{10}$ copies/swab [95%CI: -2.11 to -0.30], p=0.009), and *Parvimonas* sp. Type 1 (△=-1.03 $\log_{10}$ copies/swab [95%CI: -1.82 to -0.24], p=0.010) all decreased by ≥1 $\log_{10}$ copies/swab on average (Table 3.4, Figure 3.2C). In contrast, concentrations were more stable among DMPA-IM users, with non-significant changes (*Sneathia* spp.: △=-0.23 $\log_{10}$ copies/swab [95%CI: -0.96 to 0.49], p=0.533; *M. hominis:* △=0.29 $\log_{10}$ copies/swab [95%CI: -0.39 to 0.98], p=0.400; and *Parvimonas* sp. Type 1: △=-0.11 $\log_{10}$ copies/swab [95%CI: -0.70 to 0.49], p=0.722). The observed changes in concentration over follow-up were only statistically significantly different between contraceptive groups for *M. hominis* (difference in △=1.50 [95%CI: 0.37 to 2.63], p=0.010). In sensitivity analyses limited to women with at least one detectable bacterial taxon during follow-up, we observed the same patterns as above, however the changes were more pronounced (decrease of ≥2 $\log_{10}$ copies/swab on average) compared to the analysis that included all participants (Table 3.4, Figure S3.3).

Patterns of vaginal microbiota change among intramuscular depot-medroxyprogesterone acetate users only

In the DMPA-IM group, specimens were collected during the time when serum MPA levels are known to peak (9-14 days post-injection) [51,75]. The median number of days between enrollment and the two-week follow-up visit was 12 days (IQR: 9-22 days). There were no significant patterns in change for Nugent score or the eight bacterial taxa in the two time periods post DMPA-IM injection (Table 3.5, Figure 3.3, Figure S3.4). Total bacterial load decreased significantly in the first two weeks post-injection (enrollment to two-week visit: △=-0.64 [95%CI: -1.03 to -0.25]; p=0.001) but rebounded by the three-month follow-up visit (two-week visit to three-month visit: △=0.44 [95%CI: 0.02 to 0.86]; p=0.041).

DISCUSSION

Our study shows that vaginal bacteria previously associated with HIV acquisition were present in African women initiating postpartum contraception, although no single taxon was found in a majority of women. Findings suggested differential patterns of change in three of the eight taxa assessed, *Sneathia* spp., *M. hominis*, and *Parvimonas* sp. Type 1, among women initiating and using DMPA-IM compared to non-HC; the concentration of these three taxa decreased among women using non-HC but no change in concentration was observed among women using DMPA-IM. While significant decreases in concentrations of *Sneathia* spp., *M. hominis*, and *Parvimonas* sp. Type 1 were observed among non-HC users, there were only significantly different changes in the concentration of *M. hominis* between users of DMPA-IM and non-HC. Observed patterns in change were especially pronounced among women with these bacteria present at contraception initiation. DMPA-IM use did not increase the likelihood of detection or the concentration of these bacteria, and changes in bacterial concentrations were not correlated with the timing of peak plasma MPA levels.

There are several potential explanations for the observed stability in *Sneathia* spp., *M. hominis*, and *Parvimonas* sp. Type 1 concentrations at the three-month follow-up visit observed among DMPA-IM users. Recent studies of the vaginal microbiome in pregnancy and the postpartum period have found that the vaginal microbiota in the one to six weeks after delivery are much more diverse, with lower levels of *Lactobacillus* spp., compared to both pre-pregnancy and during pregnancy [64,65]. This time also coincides with a known period of higher HIV risk in women [80]. DMPA-IM use could stabilize concentrations of vaginal anerobic bacteria in the postpartum period by allowing a hypoestrogenic state to persist, while non-HC users returned to the more optimal vaginal bacterial communities seen in pre-pregnancy as their normal hormonal cycles return. Low levels of estrogen have been associated with glycogen suppression [23,25], and glycogen is an important substrate for *Lactobacillus* species [25-27]. Consequently, a reduction of glycogen in the female reproductive tract could maintain anaerobic, non-optimal vaginal microbiota. This explanation aligns well with our observations: DMPA-IM initiation was not associated with large increases in concentration of the seven bacterial taxa linked to HIV acquisition in our study, however it appears that DMPA-IM might support continued higher concentrations compared to non-HC. Alternatively, however, this pattern could be due to the fact that participants selected their own contraceptive method, and women who rely on non-hormonal methods alone appear to be behaviorally and demographically different from women who choose highly effective contraceptive methods [68], including in sexual and personal hygiene practices. We collected information on potential confounders to enable for control of this confounding in statistical analyses, however some measures relied on self-report and are likely imperfect.

Our results generally agree with recent findings from a study in Zimbabwe conducted among HIV-negative, nonpregnant women [38]. Achilles et al. reported that use of DMPA-IM did not significantly change the concentration of three common BV-associated species: *G. vaginalis*, *Atopobium vaginae*, and *Megasphaera*-like bacterium phylotype 1. Unlike the Zimbabwean cohort, we did observe a significant decrease in Nugent score among the DMPA-IM users in our study, which is congruent with existing literature [36]. Women who chose DMPA-IM in our study had much higher Nugent scores at enrollment compared to women who chose non-HC, however, giving them more room to decrease in Nugent score and posing a concern that these groups of women were not completely comparable. This discrepancy in Nugent score at enrollment may be explained, at least in part, by the fact that more DMPA-IM users had resumed sexual intercourse prior to enrollment.

Strengths of this study include that we enrolled participants who were free of any hormonal contraceptives prior to enrollment and received their first DMPA-IM injection from study staff. This design eliminated contamination from other contraceptives, as well as uncertainty in type of injectable hormonal contraceptive used due to self-report. It also allowed for comparison of pre- and post-DMPA-IM effects, as women could serve as their own controls. Another strength is that specimens were collected during the time period when MPA levels peak [51,75], allowing for the examination of a relationship between vaginal microbiota and high MPA plasma levels. Furthermore, our study focused on African women, the geographical population at highest risk of STI/HIV acquisition from DMPA-IM use. Lastly, we used qPCR to measure the absolute abundance of vaginal bacteria, rather than relying on relative abundance measures.

Our study was limited in power because most women had concentrations below the lower limit of detection for each taxon, and most women remained in the same detection category in which they started; this restricted our analytic options and some models contained relatively few women. This limitation exemplifies the complexities of working with qPCR data. Additionally, we had a modest loss to follow up (19% of women), further reducing our power to detect true differences. A second limitation is that participants selected their own contraceptive method, and women choosing non-HC may not represent a counterfactual population for women who choose DMPA-IM. Also, women who chose non-HC did not return for a visit 9-14 days after enrollment since their vaginal microbiota profiles were not expected to change in the two weeks following enrollment, resulting in lack of comparable data at this time point. Our study may also be limited in generalizability due to the source population of our participants. Postpartum, amenorrheic women were enrolled to reduce the variability of the vaginal microbiome due to the menstrual cycle [24,28,67], however vaginal microbiota in the postpartum period may differ from vaginal microbiota found in menstruating women of childbearing age [64,65]. Lastly, the use of a qPCR, a targeted approach, might have missed changes in important bacterial taxa not assessed, including taxa associated with an optimal vaginal microbiota. This is an area that future work might address.

Interestingly, there was a small but notable subset of DMPA-IM users that did not follow overall trends and maintained or increased in Nugent score and concentration of *G. vaginalis*.

We looked within this subgroup for commonalities, and while we could not find any demographic, behavioral, or physical similarities, including changes in vaginal bleeding patterns and perturbation of the regular menstrual cycle, among these women from the data we collected, there is potentially a subgroup of women for which DMPA-IM has a different effect. While small, this subgroup could contribute to the increased risk of HIV seen in observational studies among DMPA-IM users. Research now suggests that there are likely at least four sub-strains/genotypes of *G. vaginalis*, not all of which produce sialidase, and which may have independent virulence factors contributing to symptoms and/or sequelae of vaginosis [81-83]. One possible area for future investigation is the relationship between DMPA-IM and sialidase-producing sub-strains of *G. vaginalis* to see if DMPA-IM has a differential effect on more pathogenic *G. vaginalis*.

Overall, these findings suggest that bacteria associated with increased HIV acquisition risk are present among postpartum women at average risk for HIV acquisition in Kenya, and that DMPA-IM use, compared to non-HC user, showed differential patterns in concentration change with three taxa associated with HIV acquisition. Our findings indicate that DMPA-IM use may support an anaerobic vaginal microbiota associated with HIV acquisition, especially among a subset of women with these species already present. High alpha- and beta-diversity may also remain for a longer period in postpartum DMPA-IM initiators, a question we plan to explore in broad range microbiome data from this cohort. While our numbers were small and we did not have as much power as we planned, these findings suggest postpartum use of DMPA-IM deserves further attention; DMPA-IM use in the postpartum period may pose different risks than DMPA-IM use in non-postpartum women of reproductive age, and more research on the effect of contraception initiation on the vaginal microbiome in this subgroup of women is warranted. While the ECHO trial has relieved apprehension about use of DMPA-IM in women in high HIV prevalence settings, specific subpopulations may still be at increased risk of HIV; DMPA-IM use in the postpartum period may pose a different risk than it does in the general population, due to high vaginal microbial diversity seen in this period. Planned microbiome studies from our cohort and among women in the ECHO trial may help to further answer questions about how DMPA-IM influences vaginal bacteria.

TABLES AND FIGURES

Table 3.1. Participant sociodemographic characteristics by contraceptive group (N=54).

Characteristic[a]	**Non-HC (n=21)**	**DMPA-IM (n=33)**
Age	26.0 (6.0)	23.1 (3.1)
Married	14 (66.7%)	32 (97.0%)
Education level at enrollment		
Primary school or less	9 (42.9%)	18 (54.6%)
Some secondary school/Completed Secondary	10 (47.6%)	10 (30.3%)
Some college/Higher education	2 (9.5%)	5 (15.2%)
Number of household residents	4.8 (1.6)	3.7 (1.1)
Median monthly household income in USD (IQR)[b]	98 (73-147)	98 (29-196)
Age at first intercourse	18.5 (2.1)	18.4 (2.3)
Lifetime number of sex partners	1.7 (1.0)	2.0 (1.1)
Number of pregnancies	2.1 (1.4)	1.8 (1.0)
Vaginal delivery (most recent delivery)	18 (85.7%)	28 (84.8%)
Time since most recent delivery (weeks)	9.9 (3.5)	7.1 (2.0)
Previous use of any family planning method	10 (47.6%)	11 (33.3%)

Abbreviations: DMPA-IM, intramuscular depot-medroxyprogesterone acetate; HC, hormonal contraception; IQR, interquartile range; SD, standard deviation.

[a] Data presented as n (%) or mean (SD) unless otherwise noted.

[b] Non-HC users: n=16; DMPA-IM users: n=19.

Table 3.2. Key participant characteristics by contraceptive group at enrollment and the three-month follow-up visit.

	Enrollment (N=54)		**Follow-up (N=44)**	
Characteristic[a]	**Non-HC (n=21)**	**DMPA-IM (n=33)**	**Non-HC (n=18)**	**DMPA-IM (n=26)**
Current breastfeeding	21 (100%)	33 (100%)	18 (100%)	26 (100%)
Exclusive breastfeeding	20 (95.2%)	31 (93.9%)	13 (72.2%)	19 (73.1%)
Mean vaginal pH[b]	5.3 (0.9)	5.5 (0.6)	4.9 (0.8)	5.2 (0.8)
Nugent-BV (Nugent score ≥7)	7 (33.3%)	19 (57.6%)	5 (27.8%)	8 (30.8%)
Mean Nugent score	3.6 (3.7)	5.3 (3.8)	2.7 (3.7)	3.3 (3.7)
Self-reported abnormal vaginal discharge[1]	3 (14.3%)	0	0	2 (7.7%)
Clinician-reported abnormal vaginal discharge[c]	0	2 (6.1%)	0	1(3.9%)
Abnormal vaginal itching[1,d]	5 (23.8%)	0	1 (5.6%)	3 (11.5%)
Vaginal washing[1]	13 (61.9%)	17 (51.5%)	13 (72.2%)	22 (84.6%)
Vaginal bleeding[1]	3 (14.3%)	9 (27.3%)	7 (38.9%)	17 (65.4%)
Use of antibiotics or metronidazole[2]	0	1 (3.0%)	1 (5.6%)	2 (7.7%)
Resumed sexual intercourse since delivery (enrollment) or recent sex[1] (follow-up)	6 (28.6%)	17 (51.5%)	10 (55.6%)	18 (69.2%)
Currently using condoms[3,d]	1 (16.7%)	0	8 (80.0%)	2 (11.1%)

Abbreviations: BV, bacterial vaginosis; DMPA-IM, intramuscular depot-medroxyprogesterone acetate; HC, hormonal contraception; SD, standard deviation.

Enrollment: visit DMPA-IM was administered; Follow-up: visit three-months post DMPA-IM initiation.

[a] Data presented as n (%) or mean (SD) unless otherwise noted.

[b] Follow-up Non-HC users: n=17; Follow-up DMPA-IM users: n=23.

[c] Enrollment Non-HC users: n=20; Baseline DMPA-IM users: n=32; Follow-up Non-HC users: n=17; Follow-up DMPA-IM users: n=23.

[d] Enrollment Non-HC users: n=20; [d] Follow-up DMPA-IM users: n=17.

[1] Past week.

[2] Past month.

[3] Percentage among of those reporting recent sex.

Table 3.3. Populations of key vaginal bacteria in postpartum women at enrollment and three-months post-initiation of DMPA-IM or non-hormonal contraception.

	Detectability at enrollment[a]			**Concentration among women with taxon detected at enrollment[b,c]**		
Taxa	**Non-HC (n=21)**	**DMPA-IM (n=33)**	**P-value[d]**	**Non-HC**	**DMPA-IM**	**P-value[e]**
Total bacterial load	--	--		8.31 (0.84)	8.59 (0.58)	0.158
G. vaginalis	15 (71.4)	28 (84.9)	0.233	6.86 (1.61)	7.43 (1.69)	0.294
Sneathia spp.	11 (52.4)	15 (45.5)	0.619	6.68 (1.88)	5.69 (1.89)	0.199
Eggerthella sp. Type 1	9 (42.9)	12 (36.4)	0.633	6.58 (1.78)	7.17 (1.54)	0.431
M. hominis	9 (42.9)	7 (21.2)	0.089	6.66 (1.81)	5.10 (1.72)	0.103
G. asaccharolytica	6 (28.6)	11 (33.3)	0.713	6.95 (1.09)	5.72 (1.04)	**0.037**
Parvimonas sp. Type 1	8 (38.1)	6 (18.2)	0.104	5.77 (2.27)	4.61 (1.58)	0.308
Parvimonas sp. Type 2	7 (33.3)	11 (33.3)	1.000	6.72 (2.19)	5.93 (2.09)	0.454
Megasphaera spp. 1 & 2	3 (14.3)	5 (15.2)	1.000	6.09 (3.08)	7.53 (0.38)	0.315

	Detectability at follow-up[a]			**Concentration among women with taxon detected at follow-up[b,c]**		
Taxa	**Non-HC (n=18)**	**DMPA-IM (n=26)**	**P-value[d]**	**Non-HC**	**DMPA-IM**	**P-value[e]**
Total bacterial load	--	--		8.74 (0.50)	8.31 (0.89)	0.072
G. vaginalis	13 (72.2)	19 (73.1)	0.950	7.06 (1.52)	7.24 (1.42)	0.735
Sneathia spp.	6 (33.3)	11 (42.3)	0.548	6.14 (2.10)	5.11 (2.31)	0.379
Eggerthella sp. Type 1	6 (33.3)	10 (38.5)	0.728	6.71 (2.08)	6.03 (2.29)	0.562
M. hominis	5 (27.8)	6 (23.1)	0.723	5.94 (2.64)	5.28 (1.55)	0.620
G. asaccharolytica	3 (16.7)	8 (30.8)	0.480	6.30 (2.32)	5.05 (1.76)	0.357
Parvimonas sp. Type 1	2 (11.1)	5 (19.2)	0.682	6.77 (1.99)	3.70 (0.83)	**0.025**
Parvimonas sp. Type 2	4 (22.2)	5 (19.2)	1.000	5.70 (3.19)	6.11 (2.41)	0.831
Megasphaera spp. 1 & 2	2 (11.1)	4 (15.4)	1.000	4.21 (2.89)	5.60 (2.43)	0.564

Abbreviations: DMPA-IM, intramuscular depot-medroxyprogesterone acetate; HC, hormonal contraception; SD, standard deviation.
Enrollment: visit DMPA-IM was administered; Follow-up: visit three-months post DMPA-IM.
Bolded text indicates significant values.

[a] Data displayed as n (%).

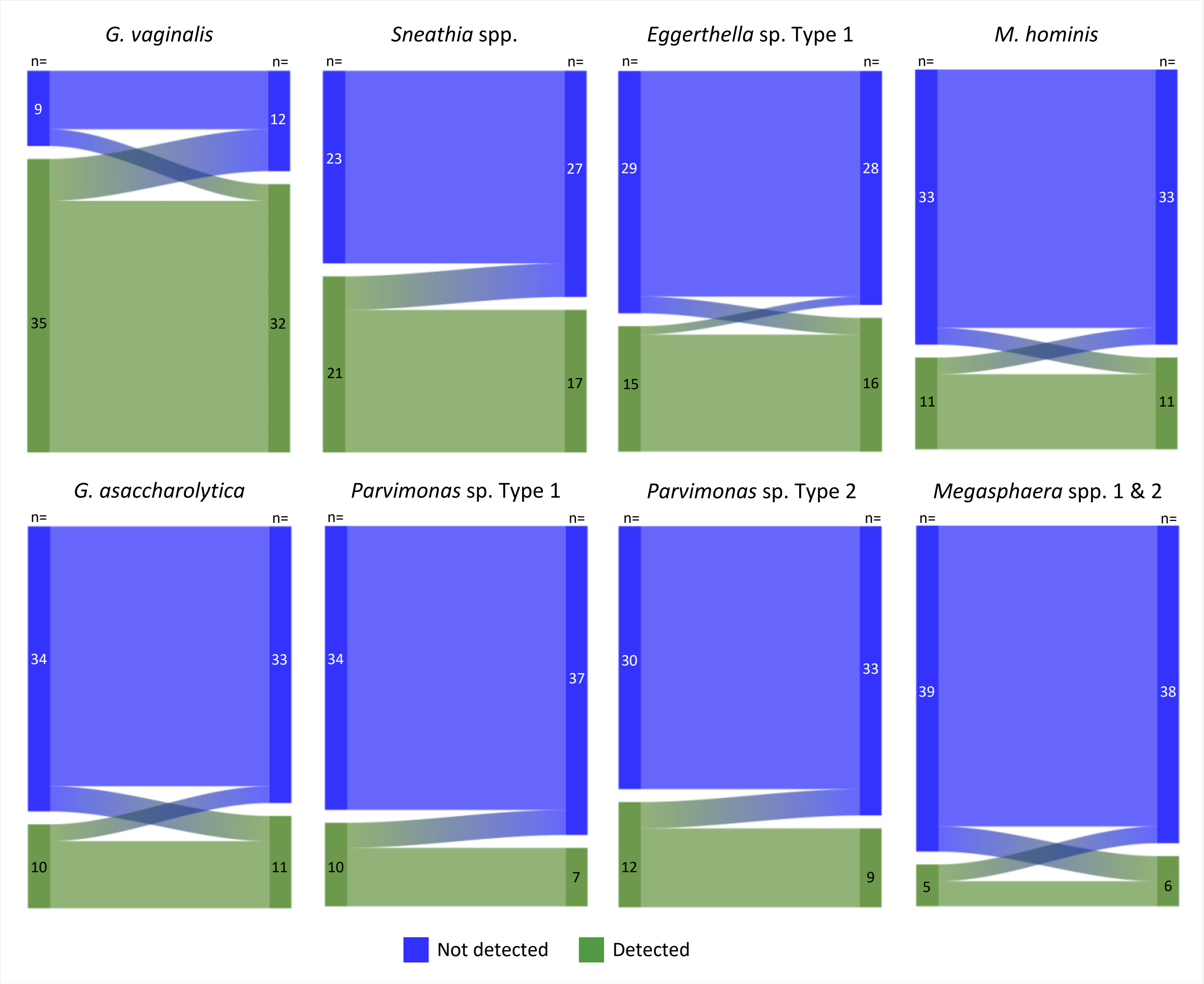
[b] Concentration displayed as $\log_{10}$ copies/swab.

[c] Data displayed as mean (SD).

[d] Chi-square test or Fisher's exact test (for comparisons containing cells with <5 observations).

[e] T-test.

Figure 3.1. Change in presence/absence of the eight bacterial taxa from enrollment to the three-month follow-up visit among the 44 women who returned for the follow-up visit.



Enrollment: visit DMPA-IM was administered; Follow-up: visit three-months post DMPA-IM.

Table 3.4. Change in Nugent score, total bacterial load, and concentration of vaginal taxa over time by contraceptive group, and difference in change between contraceptive groups during follow-up.

All participants[a]

	non-HC	**95% CI**	**P-value**	**DMPA-IM**	**95% CI**	**P-value**	**Difference**	**95% CI**	**P-value**
Nugent score	-0.73	-2.70, 1.24	0.469	-1.90	-3.38, -0.41	**0.012**	-1.17	-3.64, 1.30	0.354
Total 16S rRNA	0.33	-0.13, 0.79	0.160	-0.25	-0.60, 0.10	0.159	-0.58	-1.16, -0.01	**0.049**
G. vaginalis	0.22	-1.02, 1.46	0.725	-0.87	-1.81, 0.06	0.067	-1.10	-2.65, 0.45	0.166
Sneathia spp.	-1.09	-2.05, -0.13	**0.026**	-0.23	-0.96, 0.49	0.533	0.86	-0.34, 2.06	0.161
Eggerthella sp. Type 1	-0.12	-1.18, 0.94	0.827	-0.06	-0.87, 0.74	0.875	0.05	-1.28, 1.38	0.936
M. hominis	-1.21	-2.11, -0.3	**0.009**	0.29	-0.39, 0.98	0.400	1.50	0.37, 2.63	**0.010**
G. asaccharolytica	-0.63	-1.53, 0.28	0.174	0.06	-0.62, 0.75	0.855	0.69	-0.44, 1.82	0.233
Parvimonas sp. Type 1	-1.03	-1.82, -0.24	**0.010**	-0.11	-0.70, 0.49	0.722	0.92	-0.07, 1.91	0.067
Parvimonas sp. Type 2	-0.80	-1.79, 0.18	0.110	-0.26	-1.03, 0.50	0.501	0.54	-0.71, 1.79	0.397
Megasphaera spp. 1 & 2	-0.13	-0.99, 0.72	0.758	-0.15	-0.80, 0.49	0.640	-0.02	-1.09, 1.05	0.972

Participants with ≥1 detectable value during follow-up[b]

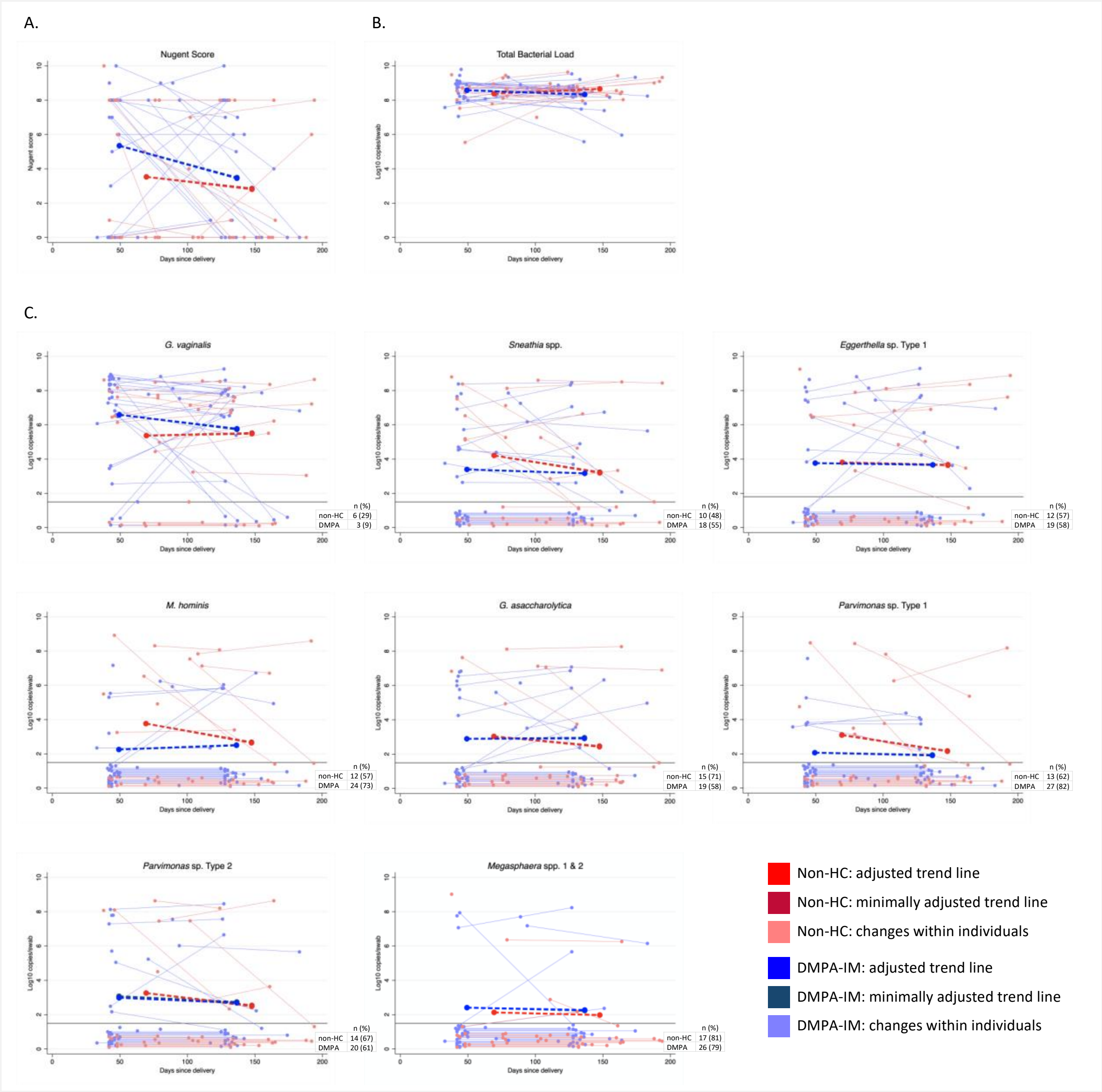
	n	non-HC	95% CI	P-value	DMPA-IM	95% CI	P-value	Difference	95% CI	P-value
G. vaginalis	47	0.39	-1.30, 2.09	0.647	-1.01	-2.12, 0.10	0.075	-1.40	-3.42, 0.62	0.174
Sneathia spp.	26	-2.19	-3.74, -0.63	**0.006**	-0.62	-1.85, 0.62	0.327	1.57	-0.41, 3.56	0.121
Eggerthella sp. Type 1	23	-0.27	-2.36, 1.82	0.801	-0.30	-1.94, 1.33	0.716	-0.03	-2.69, 2.62	0.980
M. hominis	18	-3.06	-4.99, -1.13	**0.002**	1.24	-0.63, 3.11	0.194	4.30	1.61, 6.99	**0.002**
G. asaccharolytica	20	-1.96	-4.18, 0.26	0.084	0.14	-1.30, 1.58	0.848	2.10	-0.55, 4.74	0.120
Parvimonas sp. Type 1	14	-2.50	-4.63, -0.36	**0.022**	-0.92	-3.21, 1.38	0.435	1.58	-1.56, 4.72	0.323
Parvimonas sp. Type 2	20	-2.71	-5.35, -0.07	**0.045**	-0.56	-2.41, 1.29	0.552	2.14	-1.08, 5.37	0.193
Megasphaera spp. 1 & 2	11	-0.91	-4.75, 2.94	0.644	-0.80	-3.31, 1.71	0.535	0.11	-4.48, 4.71	0.962

Abbreviations: CI, confidence interval; DMPA-IM, intramuscular depot-medroxyprogesterone acetate; HC, hormonal contraception.
Enrollment: visit DMPA-IM was administered; Follow-up: visit three-months post DMPA-IM.
Concentration displayed as $\log_{10}$ copies/swab. All values below LLD were set to half the LLD for that assay.
Bolded text indicates significant values.

[a] Mean change for each contraceptive group and difference in change estimated using GEE with an interaction term between contraceptive group and days from enrollment to vaginal swab collection and adjusted for days from delivery to enrollment, age, marital status, and resumption of intercourse at enrollment.

[b] Mean change for each contraceptive group and difference in change among women with ≥1 detectable value for that taxon during follow-up estimated using GEE with an interaction term between contraceptive group and days from enrollment to vaginal swab collection and adjusted for days from delivery to enrollment.

Figure 3.2. (A) Nugent score, (B) total bacterial load, and (C) concentration of vaginal taxa over time, by contraceptive group, with fitted trend lines for mean change among all women.



Abbreviations: DMPA-IM, intramuscular depot-medroxyprogesterone acetate; HC, hormonal contraception.

Bacterial concentrations were $\log_{10}$ transformed to normalize their distribution.

All values below LLD (black horizontal bar on graph) were equal, but values were jittered to allow for visualization of all observations.

Trend lines for mean change within each contraceptive group were estimated using GEE with an interaction term between contraceptive group and days from delivery to enrollment and adjusted for days from enrollment to vaginal swab collection; trend lines show mean enrollment and exit dates for each group.

Table 3.5. Change in Nugent score, total bacterial load, and concentration of vaginal taxa over time among DMPA-IM users only, and difference in change between time periods.

All participants[a]

	Enrollment to two-week visit	95% CI	P-value	Two-week to three-month visit	95% CI	P-value	Difference	95% CI	P-value
Nugent score	-1.31	-2.74, 0.12	0.073	-0.62	-2.17, 0.93	0.433	0.69	-1.90, 3.27	0.601
Total 16S rRNA	-0.64	-1.03, -0.25	**0.001**	0.44	0.02, 0.86	**0.041**	1.08	0.37, 1.78	**0.003**
G. vaginalis	-0.66	-1.55, 0.23	0.146	-0.22	-1.18, 0.74	0.658	0.44	-1.16, 2.04	0.589
Sneathia spp.	-0.09	-0.67, 0.49	0.761	-0.06	-0.70, 0.57	0.846	0.03	-1.03, 1.08	0.959
Eggerthella sp. Type 1	0.01	-0.75, 0.77	0.974	-0.06	-0.88, 0.76	0.882	-0.07	-1.44, 1.29	0.914
M. hominis	0.03	-0.49, 0.55	0.915	0.33	-0.23, 0.90	0.248	0.30	-0.63, 1.24	0.526
G. asaccharolytica	-0.19	-0.76, 0.37	0.500	0.37	-0.24, 0.98	0.236	0.57	-0.46, 1.59	0.278
Parvimonas sp. Type 1	0.06	-0.28, 0.39	0.741	-0.19	-0.55, 0.18	0.312	-0.24	-0.85, 0.36	0.431
Parvimonas sp. Type 2	0.15	-0.49, 0.79	0.647	-0.40	-1.12, 0.32	0.275	-0.55	-1.74, 0.64	0.364
Megasphaera spp. 1 & 2	-0.23	-0.73, 0.26	0.357	0.16	-0.37, 0.70	0.550	0.40	-0.50, 1.29	0.384

Participants with ≥1 detectable value during follow-up[b]

	n	Enrollment to two-week visit	95% CI	P-value	Two-week to three-month visit	95% CI	P-value	Difference	95% CI	P-value
G. vaginalis	31	-0.68	-1.65, 0.29	0.167	-0.24	-1.28, 0.81	0.658	0.45	-1.30, 2.19	0.617
Sneathia spp.	16	-0.23	-1.22, 0.76	0.651	0.13	-1.05, 1.32	0.825	0.36	-1.48, 2.21	0.700
Eggerthella sp. Type 1	16	-0.08	-1.54, 1.37	0.910	-0.27	-2.04, 1.51	0.769	-0.18	-2.93, 2.57	0.897
M. hominis	10	0.05	-1.34, 1.44	0.943	1.48	-0.09, 3.04	0.064	1.43	-1.15, 4.00	0.278
G. asaccharolytica	14	-0.31	-1.12, 0.50	0.456	1.39	0.46, 2.32	**0.003**	1.70	0.22, 3.19	**0.025**
Parvimonas sp. Type 1[c]	7	--			--			--		

Parvimonas sp. Type 2	17	0.54	-0.71, 1.79	0.397	-1.19	-2.70, 0.32	0.124	-1.73	-4.13, 0.68	0.159
Megasphaera spp. 1 & 2[c]	8	--			--			--		

Abbreviations: CI, confidence interval; DMPA-IM, intramuscular depot-medroxyprogesterone acetate.

Enrollment: visit DMPA-IM was administered; two-week visit: visit 14 days post DMPA-IM; Follow-up: visit three-months post DMPA-IM.

Concentration displayed as $\log_{10}$ copies/swab. All values below LLD were set to half the LLD for that assay.

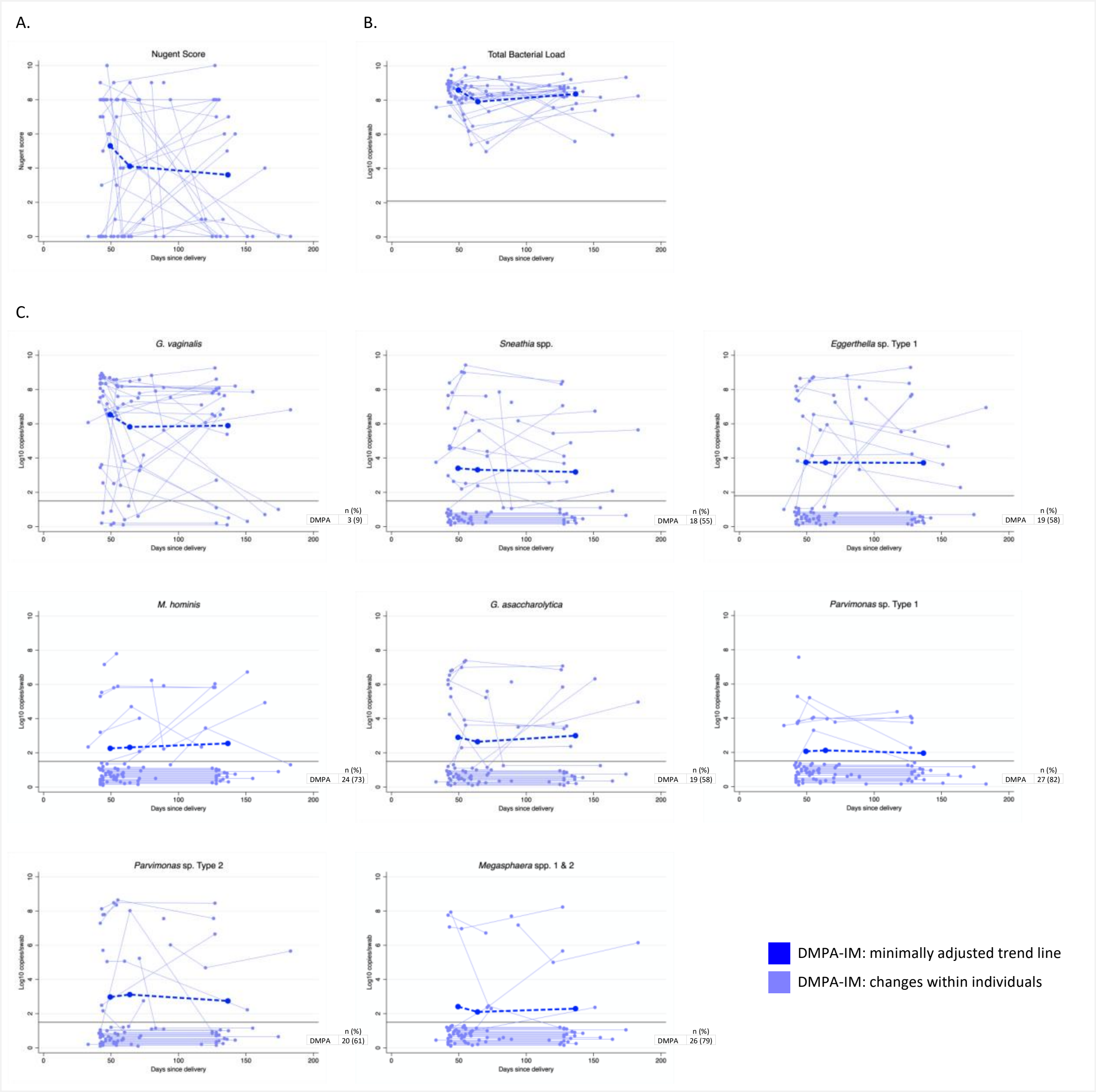
Bolded text indicates significant values.

[a] Mean change in the two time periods and difference in change between time periods estimated using GEE adjusted for days from enrollment to vaginal swab collection and days from delivery to enrollment.

[b] Mean change in the two time periods and difference in change between time periods among women with ≥1 detectable value for that taxon during follow-up estimated using GEE adjusted for days from enrollment to vaginal swab collection and days from delivery to enrollment.

[c] Too few observations to model data.

Figure 3.3. (A) Nugent score, (B) total bacterial load, and (C) concentration of vaginal taxa over time among DMPA-IM users only with fitted trend lines for mean change among all women.



Abbreviations: DMPA-IM, intramuscular depot-medroxyprogesterone acetate.

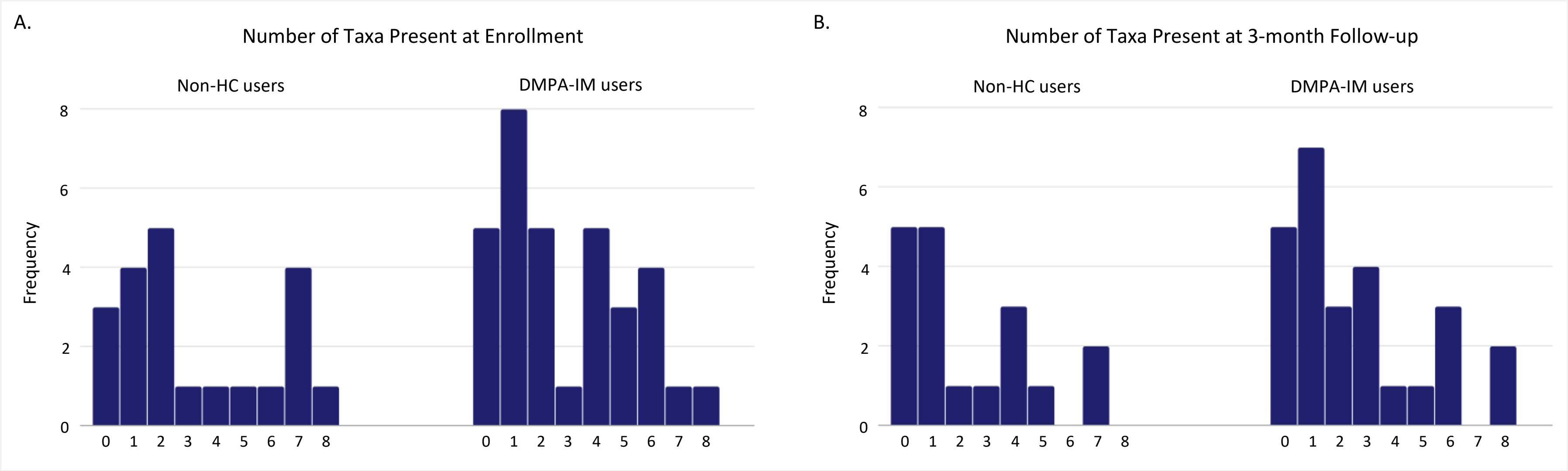
Bacterial concentrations were $\log_{10}$ transformed to normalize their distribution.

All values below LLD (black horizontal bar on graph) were equal, but values were jittered to allow for visualization of all observations.

Trend lines for mean change were estimated using GEE adjusted for days from delivery to enrollment and days from enrollment to vaginal swab collection; trend lines show mean enrollment and exit dates.

SUPPLEMENTAL TABLES AND FIGURES

Figure S3.1. Histogram of the number of taxa present (out of the eight assessed) at (A) Enrollment (n=54) and (B) the three-month follow-up visit (n=44).



Abbreviations: DMPA-IM, intramuscular depot-medroxyprogesterone acetate; HC, hormonal contraception.
Enrollment: visit DMPA-IM was administered; Follow-up: visit three-months post DMPA-IM.

Table S3.1. McNemar's test for change in Nugent-BV category and detectability of bacterial taxa from enrollment to three-month follow-up by contraceptive group.

	Group	P-value[a]
Nugent-BV category (Nugent score of 4-10 vs. 0-3)	non-HC	0.63
	DMPA-IM	0.34
G. vaginalis	non-HC	-- [b]
	DMPA-IM	0.45
Sneathia spp.	non-HC	0.25
	DMPA-IM	1.00
Eggerthella sp. Type 1	non-HC	1.00
	DMPA-IM	0.50
M. hominis	non-HC	0.50
	DMPA-IM	0.50
G. asaccharolytica	non-HC	1.00
	DMPA-IM	0.63
Parvimonas sp. Type 1	non-HC	0.25
	DMPA-IM	-- [b]
Parvimonas sp. Type 2	non-HC	1.00
	DMPA-IM	0.50
Megasphaera spp. 1 & 2	non-HC	1.00
	DMPA-IM	1.00

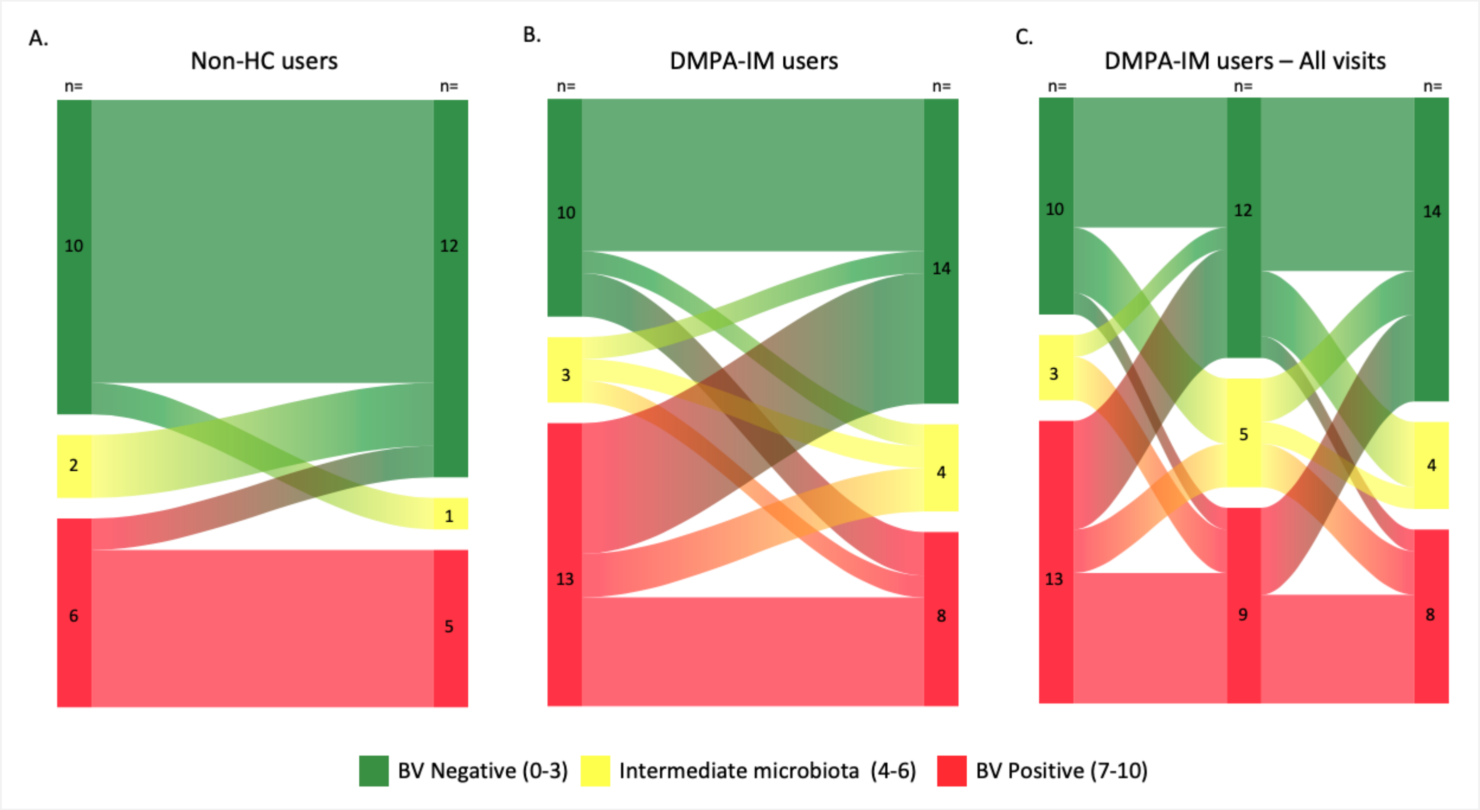
Abbreviations: BV, bacterial vaginosis; DMPA-IM, intramuscular depot-medroxyprogesterone acetate; HC, hormonal contraception.

Enrollment: visit DMPA-IM was administered; Follow-up: visit three-months post DMPA-IM.

[a] Exact McNemar significance probability.

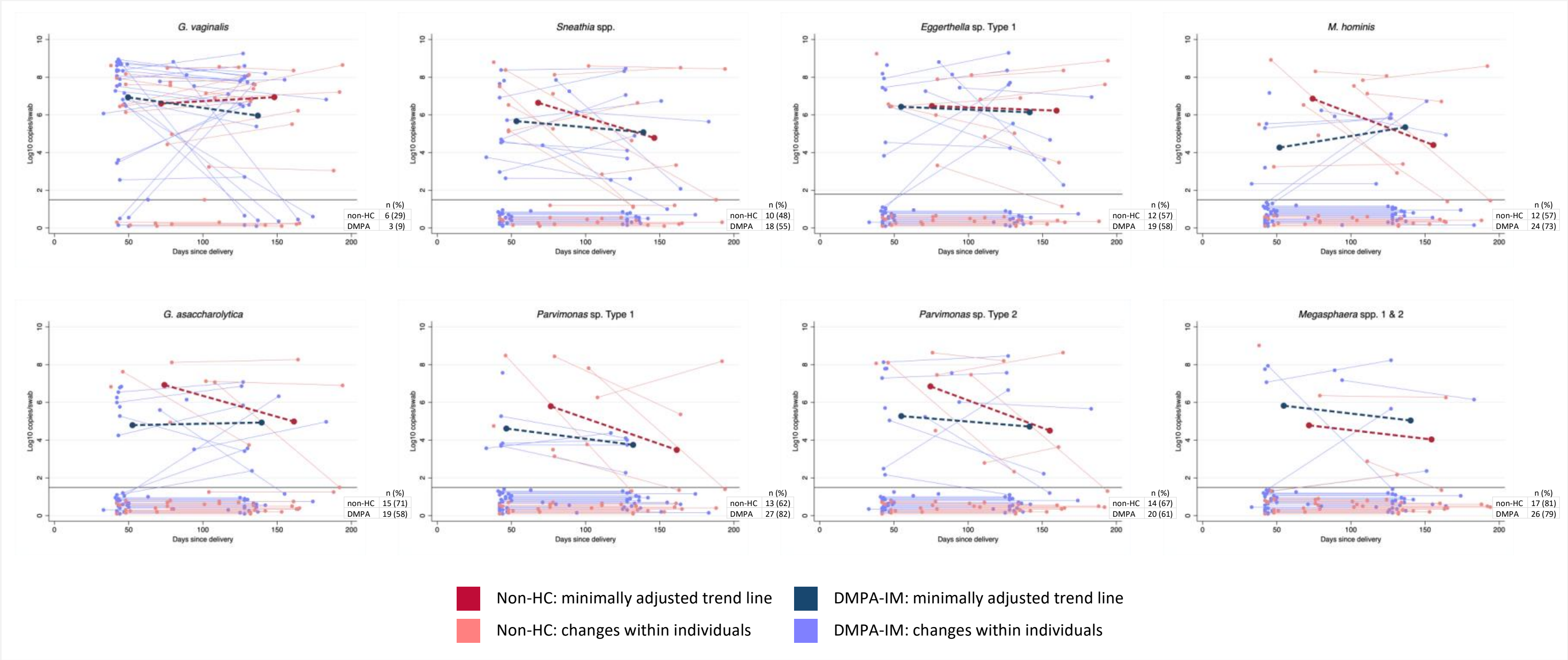
[b] No switches in detectability.

Figure S3.2. Change in Nugent-BV category from enrollment to the three-month follow-up for (A) non-HC users (n=18) and (B) DMPA-IM users (n=26), among the 44 women who returned for follow-up. (C) Change in Nugent-BV category from enrollment to the two-week post-injection visit to the three-month follow-up visit among DMPA-IM users (n=26) who attended every visit.



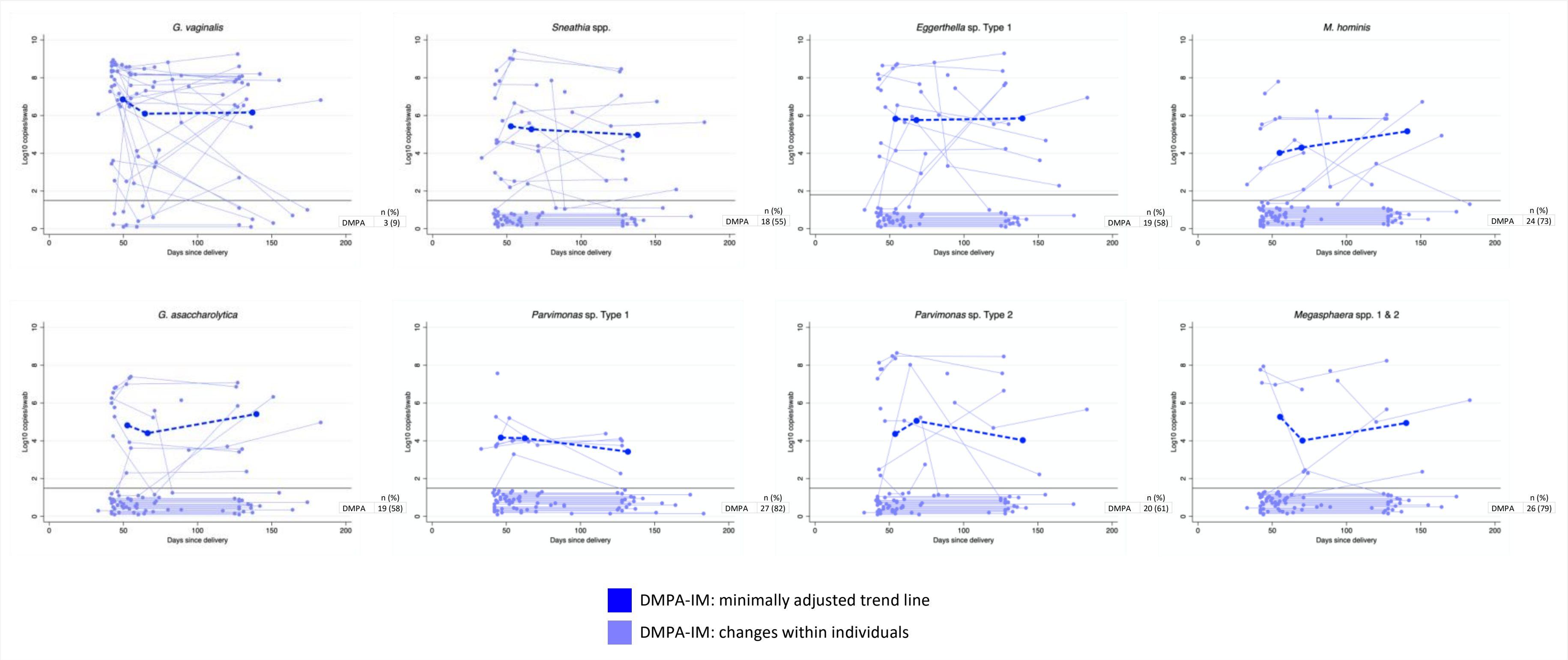
Abbreviations: BV, bacterial vaginosis; DMPA-IM, intramuscular depot-medroxyprogesterone acetate; HC, hormonal contraception.
Enrollment: visit DMPA-IM was administered; two-week visit: visit 14 days post DMPA-IM; Follow-up: visit three-months post DMPA-IM.

Figure S3.3. (A) Nugent score, (B) total bacterial load, and (C) concentration of vaginal taxa over time, by contraceptive group, with fitted trend lines for mean change among women ≥1 detectable value during follow-up.



Abbreviations: DMPA-IM, intramuscular depot-medroxyprogesterone acetate; HC, hormonal contraception.
Bacterial concentrations were $\log_{10}$ transformed to normalize their distribution.
All values below LLD (black horizontal bar on graph) were equal, but values were jittered to allow for visualization of all observations.
Trend lines for mean change in concentration among women ≥1 detectable value during follow-up within each contraceptive group were estimated using GEE with an interaction term between contraceptive group and days from delivery to enrollment and adjusted for days from enrollment to vaginal swab collection; trend lines show mean enrollment and exit dates for each group.

Figure S3.4. (A) Nugent score, (B) total bacterial load, and (C) concentration of vaginal taxa over time among DMPA-IM users only with fitted trend lines for mean change among women ≥1 detectable value during follow-up.



Abbreviations: DMPA-IM, intramuscular depot-medroxyprogesterone acetate.
Bacterial concentrations were $\log_{10}$ transformed to normalize their distribution.
All values below LLD (black horizontal bar on graph) were equal, but values were jittered to allow for visualization of all observations.
Trend lines for mean change in concentration among women ≥1 detectable value during follow-up were estimated using GEE adjusted for days from delivery to enrollment and days from enrollment to vaginal swab collection; trend lines show mean enrollment and exit dates.

CHAPTER 4. Comparison of Different Polymerase Chain Reaction Measures of Vaginal Microbiota for the Prediction of HIV Acquisition

ABSTRACT

Background: The vaginal microbiome is commonly evaluated using broad-range polymerase chain reaction (PCR) with next generation sequencing (NGS) or taxon-specific quantitative PCR (qPCR); both approaches have advantages and disadvantages. Assessing the predictive value of NGS vs. qPCR for HIV acquisition may help identify optimal PCR approach(s).

Methods: Using data from a case-control study evaluating the association between vaginal bacteria and HIV acquisition, we generated a series of models using logistic regression to assess the performance of bacterial concentration and relative abundance measures as predictors of HIV. Vaginal swabs were analyzed using broad-range PCR with pyrosequencing using the Roche 454 pyrosequencer platform (NGS) and 20 selected taxa via taxon-specific qPCR. We performed principal components analysis (PCA) to assess the combined predictive value of bacterial abundances of taxa of interest. We also performed least absolute shrinkage and selection operator (LASSO) to select the most informative taxa for HIV prediction. Level of discrimination was quantified using the area under the receiver-operating curve (AUC) and AUCs were compared between models.

Results: Vaginal swabs from 110 women were included in analyses. Median concentrations and relative abundances of bacteria associated with HIV acquisition were low. The reference model containing demographic and risk factor variables had an acceptable discrimination value (AUC=0.706, 95% Confidence Interval (CI): 0.609-0.804). The first three principal components (PCs) from qPCR data resulted in an AUC=0.787 (95%CI: 0.702-0.872) and the first three PCs

from NGS data resulted in an AUC=0.761 (95%CI: 0.672-0.850). AUCs from LASSO selection resulted in similar discrimination (AUC=0.771 vs. AUC=0.784, respectively).

Conclusion: We found the abundance of bacterial taxa associated with HIV acquisition from qPCR, in the form of concentration, and from broad-range PCR with NGS, in the form of relative abundance, provide similar predictive discrimination between women who acquired HIV and women who remained HIV uninfected.

INTRODUCTION

The understanding of the human microbiome has developed markedly in the last three decades since culture-independent methods to take bacterial census became available.[84] The vaginal microbiome is commonly evaluated using broad-range polymerase chain reaction (PCR) with next generation sequencing (NGS), hereafter referred to as NGS, or taxon-specific quantitative PCR (qPCR). Both approaches have advantages and disadvantages.

NGS targets conserved regions of the 16S rRNA gene and is advantageous because it provides a comprehensive picture of the microbiome in a sample and does not require the *a priori* determination of which bacterial taxa will be assessed. Advances in NGS have allowed sampling of bacterial DNA at new depths, providing a more detailed picture of the vaginal microbiome.[45] NGS is useful for assessing community composition and bacterial diversity, but it only provides the relative abundance of bacterial taxa present in a specimen.[45] Additionally, analytic decisions at all stages of processing, including the operational taxonomic unit (OTU) generation strategy, choice of reference database, clustering algorithm, and software package, can greatly affect results.[46,47] In contrast, qPCR measures the absolute concentration of pre-defined bacterial taxa through the use of taxon-specific primers. Data describing taxon-specific concentrations enables the evaluation of changes in bacterial concentration over time, which may be inapparent from relative abundance measures. However, qPCR does not provide an overall view of vaginal microbiota, the major limitation of this approach.[45]

Using both PCR approaches together provides complementary data. To our knowledge, no published studies have directly compared the predictive value of different measures of the vaginal microbiota for biological outcomes such as HIV acquisition. Multiple studies using molecular methods have identified vaginal bacteria associated with increased HIV acquisition risk. *Eggerthella* species Type 1,[17] *Gemella asaccharolytica*,[17] *Leptotrichia/Sneathia* species,[16,17] *Megasphaera* species,[17] *Mycoplasma* species,[16,17] *Parvimonas* species type 2,[17] and *Prevotella* species[16] have been associated with HIV acquisition, most in a dose-dependent matter. In addition to the peer-reviewed literature, two abstracts presented at major HIV conferences reported an association between *P. bivia* and increased HIV acquisition risk.[73,74] Additionally, *Eggerthella* sp. Type 1, *G. asaccharolytica*, Leptotrichia*/Sneathia* spp., *Megasphaera* sp. Type 1, *M.* hominis, and *Parvimonas* sp. type 2 were found to be associated with increased HIV acquisition risk,[74] replicating previous findings.[17]

Assessing the predictive value of NGS vs. qPCR for HIV acquisition may help identify optimal PCR approach(s) for use in future research and clinical applications. We hypothesized that the absolute concentrations of bacteria from qPCR would provide better discrimination for predicting HIV acquisition than relative abundance or overall diversity measures from NGS.

METHODS

Study Design/Setting/Subjects

This secondary analysis utilized data from a case-control study evaluating the association between specific vaginal bacteria and the risk of HIV acquisition in diverse cohorts of African women.[17] A subset of 55 cases and 55 matched controls from three of the five cohorts in the original study had both NGS and qPCR performed on their samples and were utilized for this analysis. The Mombasa Cohort included women ≥16 years old who were HIV-seronegative and self-identified as engaging in transactional sex.[85] The Mama Salama Study included women ≥14 years old, pregnant, and HIV-negative either at enrollment or documented during routine antenatal care within three months of enrollment in rural western Kenya.[86] The Partners PrEP Study included women ≥18 years old in HIV-serodiscordant heterosexual relationships in Kenya and Uganda.[87] Detailed procedures and characteristics for each cohort have been published elsewhere.[85,86,88]

The Mombasa Cohort and The Mama Salama Study used incidence density sampling to select controls. For each case, a women from the same cohort who remained HIV-negative was randomly selected and matched on calendar time to the case. Eligible controls were selected from among all HIV-negative participants who contributed a follow-up visit and had a vaginal swab for microbiota analyses taken within 28 days of the case's exposure swab date. The Partners PrEP Study selected controls at random from a pool of non-seroconverting women that was frequency-matched by site of enrollment.[17]

Measures/Data Collection/Procedures

Each research protocol received approval from both country-specific and investigator-affiliated ethical review boards. Participants provided written informed consent. Demographic, medical, and sexual history data were collected across all studies at enrollment. Participants were asked to return every 1-3 months for follow-up HIV testing. For cases, pre-HIV-infection genital samples were selected from visits at which participants were both HIV antibody negative and HIV RNA negative. All swabs for DNA extraction and bacterial PCR were stored at -80ºC and transported on dry ice to the Fred Hutchinson Cancer Research Center in Seattle, WA for analysis. The vaginal swab specimens were first analyzed using broad-range PCR with pyrosequencing using the Roche 454 pyrosequencer platform (Roche Diagnostics, Inc., Branford, CT). From exploratory analyses using the broad-range data, 20 taxa were selected for analysis via taxon-specific qPCR.[17] For the analysis presented here, evaluation focused on taxa that have previously been associated with HIV acquisition, including *Egerthella* species Type 1, *Gemella asaccharolytica, Lactobacillus crispatus*, *Leptotrichia/Sneathia* species, *Megasphaera* species Types 1 and 2*, Mycoplasma hominis*, *Parvimonas* species Type 1, *Parvimonas* species Type 2, and *Prevotella* species.[17] Where NGS and qPCR data did not give identical taxonomic classification, the genus level for that taxa was used.

Statistical Analyses

To compare differences in the output values generated from NGS (relative abundance) and qPCR (concentration), we computed Spearman correlation values for pairwise comparisons between taxa for each method. Additionally, we compared agreement between qPCR and NGS for the detectability of each taxon using Cohen's kappa.

Next, we generated a series of models using logistic regression to assess the performance of bacterial concentration and relative abundance measures as predictors of HIV acquisition. To account for matching in the data, all models were adjusted for matching factors.[89] We first assessed the predictive performance of each bacterial taxon of interest alone by fitting separate models with concentration data and relative abundance data. Next, we assessed the predictive performance of a reference model containing only demographic variables and HIV risk factors; the reference model included matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, and recent condomless sex.

In order to reduce the dimensionality of the data, we performed principal components analysis (PCA) using all bacterial taxa of interest to assess the combined predictive value of the bacterial abundances on HIV acquisition. Separate PCA models were fit for qPCR and NGS data. The first three principal components (PCs) from each PCA were then included as the predictors in logistic regression models. Two models were fit for each PCR method: 1. Matching factors plus the first three PCs only, and 2. Matching factors, the first three PCs, and covariates included in the reference model. The latter model was employed to assess if microbiome data adds additional prediction value above questionnaire data alone. Model agreement between the reference model and the four (2 PCR x 2 models) PCA prediction models were evaluated using Cohen's kappa.

Lastly, we performed least absolute shrinkage and selection operator (LASSO)[90,91] to select the most informative covariates for HIV prediction. Separate LASSO models were fit for qPCR and NGS data. For the selection of NGS taxa, we included data from all identified taxa from NGS in the LASSO model, not just the taxa of interest, as a way to utilize the additional data provided by NGS. Similar to the methodology utilized above, two series of models were fit for each PCR method. First, LASSO was performed using only selected bacterial taxa, and then logistic regression adjusted for matching factors plus selected bacterial taxa was performed. Second, LASSO was performed using qPCR/NGS data and covariates in the reference model, and then logistic regression adjusted for matching factors plus selected variables was performed. Model agreement between the four LASSO prediction models was measured using Cohen's kappa.

Level of discrimination for all models was quantified using the area under the receiver-operating curve (AUC) or c-statistic, and AUCs were compared between models. Standard values of discrimination were used, with AUCs <0.70 considered poor discrimination, values between 0.70-0.80 considered acceptable discrimination, and values >0.80 considered excellent discrimination.[92] Family wise error rate was controlled using Hochberg's step-up procedure. All analyses were conducted using Stata version 16 (StataCorp, College Station, TX).

RESULTS

Characteristics of cohort

The analysis cohort was comprised of 110 women, 14 (13%) from The Mombasa Cohort, 40 (36%) from The Mama Salama Study, and 56 (51%) from the Partners PrEP Study (Table 1). Median age was 26 years (interquartile range (IQR): 21-31), 87 (79%) of the women were married, and 23 (21%) of the women were pregnant, 39 (36%) reported a modern contraceptive (implant, intrauterine device, depo medroxyprogesterone acetate, or oral contraception), and 48 (44%) were neither pregnant nor reported modern contraception use. The number of recent sexual partners was low overall (median=1, IQR: 1-1), and 40 (36%) women reported recent condomless sex. Approximately a quarter of women (n=28, 26%) had bacterial vaginosis (BV) at enrollment, defined as a Nugent score ≥7 (Nugent-BV).

Comparison broad-range PCR with NGS and qPCR

Median concentrations and relative abundances of bacteria associated with HIV acquisition were generally low (Table 2). Hierarchically clustered Spearman correlation analyses showed that the concentrations of the taxa of interest were more highly correlated with one another than relative abundances, demonstrating important differences in the output data generated by these two methods (Figure 1). Detection of each taxon by qPCR versus NGS was evaluated using Cohen's κ. Agreement in detection between the two methods was almost perfect for *L. crispatus* and *M. hominis* (κ=0.84 and κ=0.90, respectively), substantial for *Eggerthella* sp. Type 1, *G. asaccharolytica*, and *Parvimonas* sp. Type 2/*Parvimonas* spp. (κ range: 0.72-0.76), and moderate for remainder of bacterial taxa evaluated in this study (Table 3).

Predictive model for HIV acquisition – individual taxa

The concentration data from qPCR improved prediction of HIV acquisition over chance alone for all taxa other than *L. crispatus*. All AUC values were ≤0.70, however, indicating poor discrimination (Table 4). Relative abundance data from NGS produced similar results, with the exception that *Megasphaera spp.* no longer performed better than chance. Again, discrimination was low; *G. asaccharolytica* was the only taxon that provided acceptable discrimination (AUC=0.719; 95% Confidence Interval (CI): 0.622-0.816). Concentration data and relative abundance data did not perform significantly differently from each other when analyzing each taxon alone.

Demographics/Risk factor reference model

The reference model containing matching factors, including cohort, site, and swab date, and demographic and risk factor variables, including age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, and recent condomless sex, had an acceptable discrimination value of 0.706 (95%CI: 0.609-0.804) (Table 5). This model served as the comparator to assess the value gained from PCR methods for the prediction of HIV acquisition.

Predictive model for HIV acquisition – PCA

The first three PCs from the PCA containing concentration values from all bacterial taxa of interest accounted for 83.9% of the variance in the qPCR data and had a Kaiser–Meyer–Olkin (KMO) value of 0.914, indicating the data was highly suitable for PCA. The predictive model containing matching factors and these three PCs alone resulted in an acceptable discrimination value, similar to that of the reference model (AUC=0.709, 95%CI: 0.609-0.810) (Table 5). When demographic/risk factor data were added to this model, the predictive performance improved (AUC=0.787, 95%CI: 0.702-0.872), but still only at an acceptable discrimination.

The first three PCs from the PCA containing NGS relative abundance values from all bacterial taxa of interest accounted for 63.1% of the variance in the NGS data and had a KMO value of 0.740, indicating the data was suitable for PCA, but less so than the qPCR data. The predictive model containing matching factors and these three PCs resulted in poor discrimination (AUC=0.657, 95%CI: 0.554-0.760) (Table 5). When demographic/risk factor data were added to this model, the predictive performance improved to acceptable (AUC=0.761, 95%CI: 0.672-0.850). Neither the qPCR PCA predictive model nor the NGS PCA predictive models significantly outperformed the reference model. Additionally, none of the models performed statistically differently from one another for prediction of HIV acquisition (Table 5, Figures 2A,B).

Agreement for prediction of HIV between the reference model and all predictive models containing the first three PCs ranged from poor to moderate (κ: -0.0938 to 0.4360), with models containing PCs and reference model variables performing more similarly to the reference model (Supplemental Table 1). Agreement between the predictive models containing either the first three qPCR PCs alone or the first three NGS PCs alone was substantial (κ=0.6476), indicating that the two methods provided similar individual prediction results. A similar agreement value was observed between the predictive models containing reference model variables and either the first three qPCR PCs or the first three NGS PCs (κ=0.6360).

Predictive model for HIV acquisition – LASSO

LASSO analysis selected the concentrations of *G. asaccharolytica*, *L. crispatus*, *Megasphaera* spp. Types 1 & 2, *M. hominis*, *Parvimonas* sp. Type 1, and *Parvimonas* sp. Type 2 from the qPCR data as the most informative variables for prediction. The prediction model containing matching factors and the concentration of these selected taxa resulted in an acceptable discrimination value, similar to that of the reference model (AUC=0.730, 95%CI: 0.634-0.826) (Table 6). When demographic/risk factor data were added to the LASSO analysis, age, and pregnancy/contraceptive status were additionally selected as informative variables for prediction. The predictive performance improved (AUC=0.771, 95%CI: 0.684-0.858), but still only at an acceptable discrimination.

LASSO analysis selected the relative abundance of *Lactobacillus iners*, *Leptotrichia amnionii*, *Gemella* spp., *Eggerthella* sp. Type 1, and *Escherichia coli* from all of the NGS data as the most informative variables for prediction. The prediction model containing matching factors and the relative abundance of these selected taxa also resulted in an acceptable discrimination value (AUC=0.716, 95%CI: 0.619-0.813) (Table 6). When demographic/risk factor data were added to the LASSO analysis, only pregnancy/contraceptive status was additionally selected as informative for prediction. Similar improvement in discrimination occurred in this model (AUC=0.784, 95%CI: 0.698, 0.870). As seen with the PCA iterations of these models, models with qPCR data and NGS data did not significantly outperform the reference model, nor did they perform statistically differently from one (Table 6, Figures 2C,D).

Agreement for prediction of HIV between the reference model and all predictive models containing LASSO-selected qPCR/NGS data were slight to fair (κ: 0.0483 to 0.2857) (Supplemental Table 2). Agreement between predictive models containing either selected qPCR variables or selected NGS variables from LASSO was substantial (κ=0.6315), similar to the results from the PCA analyses. Interestingly, model agreement was only moderate between the predictive models containing either the selected qPCR and reference model variables or selected NGS and reference model variables from LASSO (κ=0.4944).

Sensitivity Analysis

As a sensitivity analysis to assess how much predictive value was derived from concentration or relative abundance data versus solely detection of these taxa by the two PCR methods, the same analyses were performed using binary detection variables for all taxa of interest. The results were very similar to those described above (Supplemental Tables 3-8; Supplemental Figure 1).

DISCUSSION

We found the abundance of bacterial taxa associated with HIV acquisition from qPCR, in the form of concentration, and from broad-range PCR with NGS, in the form of relative abundance, provide similar predictive discrimination between women who acquired HIV and women who remained HIV uninfected. Whether the bacterial taxa were assessed individually or were incorporated into composite measures, such PCs from PCA, both PCR measures performed similarly in prediction models, with no significant difference in discrimination. Additionally, none of the models incorporating PCR data outperformed the reference model, which included only demographic and risk factor covariates. When NGS data were not restricted to the nine taxa of interest, but rather models were allowed to utilize all data provided by this method, qPCR and NGS again provided similar AUC values with no significant difference in discrimination.

Interestingly, results obtained from analyses using binary detection (presence/absence) of taxa from the two PCR methods provided results very similar to results from analyses using concentration and relative abundance data. The sensitivity analysis suggested that most of predictive value is driven by the bacteria being present, although some additional prediction may be derived from abundance data. This finding makes sense in the context of the findings from the main analysis, that abundance measures from the two PCR methods provide similar predictive discrimination between women who acquired HIV and women who remained HIV uninfected, as output generated from NGS and qPCR provide substantially different values, as is evidenced from the hierarchically clustered Spearman correlation analyses.

There are few papers comparing PCR methods for microbiome research. One study of the oral microbiome, which compared broad-based microbiome identification by 16S rRNA gene sequencing and targeted characterization of microbes by custom DNA microarray (Human Oral Microbe Identification Microarray, or HOMIM),[93] found that community profiles assessed by 16S rRNA pyrosequencing and HOMIM were highly correlated at the phylum level for all taxa and at the genus level for more commonly detected taxa. However, this study did not look at prediction value of the two PCR approaches for a biologically relevant endpoint. Similar results were seen in two small studies comparing gut microbiome profiles using similar methods.[94,95] The study of the oral microbiome identification assays also found high concordance between detection of common genera between the two PCR approaches. We assessed agreement in detection of bacterial taxa associated with HIV acquisition of interest by qPCR and NGS, but we did not observe the same pattern. The lack of high concordance between the two PCR methods in our study could be due to a multitude of factors, including that the median concentrations and relative abundances of taxa of interest were generally low, we assessed concordance at the species level for several taxa, and innate limitations of NGS, including sampling depth; NGS sampling depth is limited and often cannot detect minority species in a community that may be biologically important.[45] Other research in this arena has focused on comparing different NGS technologies and platforms.[47,96]

To our knowledge, this is the first study to directly compare qPCR and broad-range PCR with NGS data from the vaginal microbiome. It is also the first study to assess predicative performance of measures of the vaginal bacterial community for predicting HIV acquisition. Additionally, this nested case-control study provided a unique dataset, with collection of vaginal microbiota samples before HIV acquisition and analysis of microbiota swabs using both qPCR and NGS. Limitations of our study include its relatively small sample size; however, this is more an issue for precision and not for directly comparing the PCR methods. The purpose of this study was not to build a prediction model for use in clinical settings, but rather a method to compare qPCR and NGS data output. While a larger sample size may reduce variance and improve precision, which could lead to statistical differences in AUCs produced by the two methods, the point estimates in our study were highly comparable and this is not a major concern.

In conclusion, we found that both the absolute concentrations and relative abundances of bacterial taxa associated with HIV acquisition perform equally well as predictors for HIV acquisition; the approaches may be similar in utility for epidemiologic investigations relating the sexually transmitted infection risk. For the purposes of understanding the microbiome and for discovery, NGS has a distinct advantage; however, it is more labor intensive and requires specialized laboratory facilities. Alternatively, qPCR is easier to deploy to non-research lab settings and may be more broadly useful in real-world scenarios. Depending on setting and purpose one method may have benefit over the other, but for measuring risk of HIV acquisition, both methods appear to be useful tools.

TABLES AND FIGURES

Table 4.1. Characteristics of participants at enrollment (N=110 women).

Characteristic	
Case	55 (50.0%)
Cohort	
Mombasa Cohort	14 (12.7%)
Mama Salama Study	40 (36.4%)
Partners PrEP Study	56 (50.9%)
Nationality	
Kenya	70 (63.6%)
Tanzania	1 (0.9%)
Uganda	39 (35.5%)
Age, median (IQR)	26 (21-31)
Years of education, median (IQR)	8 (5-10)
Married	87 (79.1%)
Pregnancy/Contraceptive Status	
Pregnant	23 (20.9%)
Implant	7 (6.4%)
Intrauterine device	4 (3.6%)
Depo medroxyprogesterone acetate	21 (19.1%)
Oral contraception	7 (6.4%)
Not pregnant, no modern contraceptive	48 (43.6%)
Number of recent sexual partners, median (IQR)[a]	1 (1-1)
Frequency of sex in past month, median (IQR)[b]	3 (1-6)
Any recent condomless sex[a]	40 (36.4%)
Vaginal washing[c]	31 (57.4%)
Pelvic examination[d]	
Abnormal vaginal discharge	17 (18.5%)
Genital ulceration	0
Cervical mucopus	3 (3.3%)
Vaginal Gram stain Nugent Score[e]	
Normal (0-3)	57 (53.8%)
Intermediate (4-6)	21 (19.8%)
Bacterial vaginosis (7-10)	28 (26.4%)
Sexually transmitted infections	
Chlamydia trachomatis[f]	2 (2.6%)
Neisseria gonorrhoeae[g]	4 (4.7%)
Trichomonas vaginalis	5 (4.6%)

Bacterial community diversity	
Shannon's diversity index, median (IQR)	0.89 (0.19-2.07)

Abbreviations: IQR, interquartile range; PrEP, Pre-exposure prophylaxis.
Data presented as n (%) unless otherwise noted.

[a] Past week for Mombasa Cohort, past month for Mama Salama Study and Partners PrEP Study.

[b] Imputed for Mombasa Cohort as the past week's frequency multiplied by four to get frequency per month.

[c] Not available for Partners PrEP Study (n=28 cases and n=28 controls).

[d] Not available for n=8 cases and n=10 controls in the Mama Salama Study.

[e] Not available for n=1 case and n=2 controls in the Partners PrEP Study and n=1 control in the Mama Salama Study.

[f] Not available for n=5 cases and n=4 controls in the Mombasa Cohort and n=10 cases and n=14 controls in the Mama Salama Study.

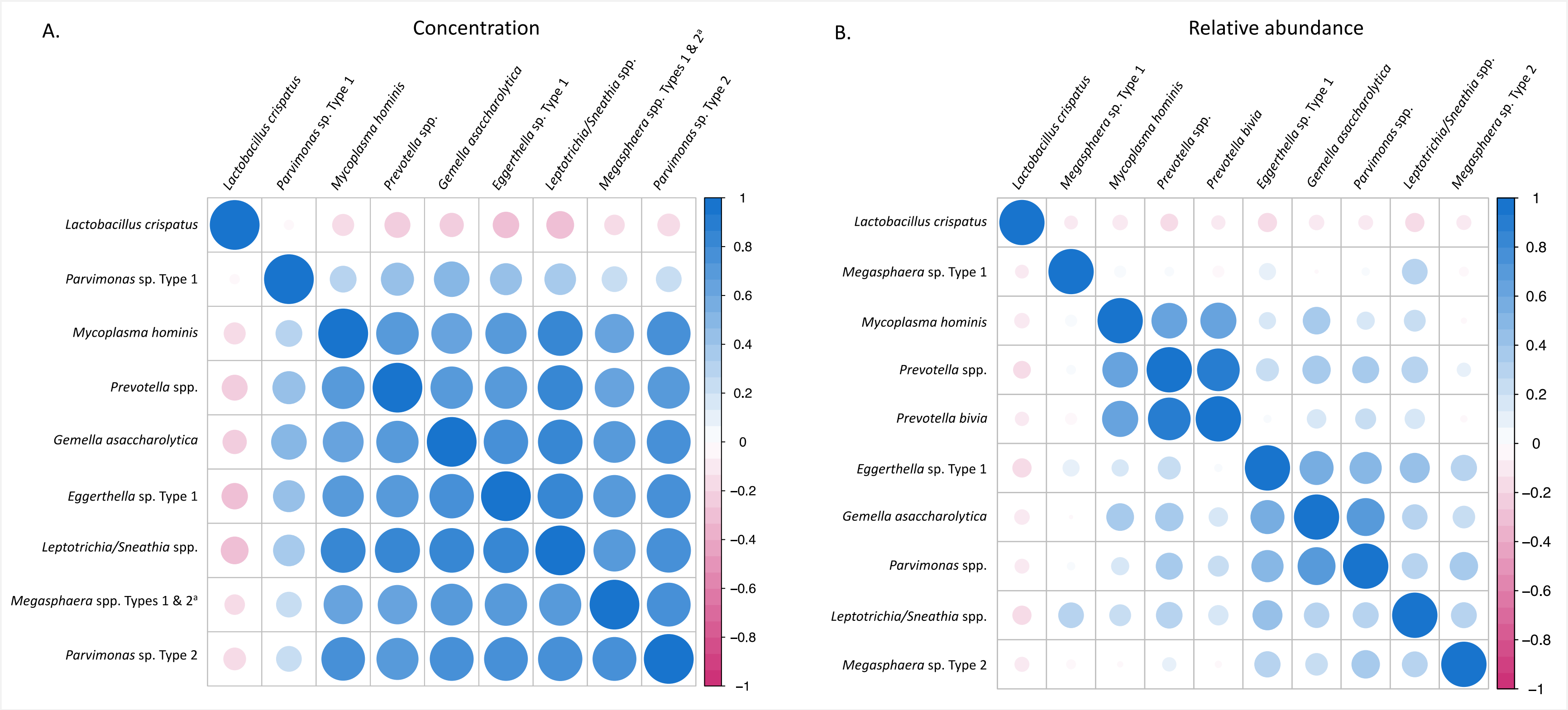
[g] Not available for n=10 cases and n=14 controls in the Mama Salama Study.

Table 4.2. Concentration and relative abundance of bacteria associated with HIV acquisition in all women (N=110).

qPCR (concentration)			**NGS (relative abundance)**		
Log$_{10}$ genome copies/swab	**Cases (n=55)**	**Controls (n=55)**	**Percent (out of 100)**	**Cases (n=55)**	**Controls (n=55)**
Eggerthella sp. Type 1	4.2 (UD-7.5)	2.3 (UD-5.1)	*Eggerthella* sp. Type 1	0 (0-0.9)	0 (0-0.02)
Gemella asaccharolytica	5.5 (UD-7.0)	UD (UD-4.3)	*Gemella asaccharolytica*	0.01 (0-0.8)	0 (0-0)
Lactobacillus crispatus	UD (UD-2.3)	UD (UD-UD)	*Lactobacillus crispatus*	0 (0-0)	0 (0-0)
Leptotrichia/Sneathia spp.	4.7 (1.9-8.4)	2.5 (UD-4.9)	*Leptotrichia/Sneathia* spp.	0.03 (0-9.0)	0 (0-0.09)
Megasphaera spp. Types 1 & 2[a]	UD (UD-7.0)	UD (UD-UD)	*Megasphaera* sp. Type 1	0 (0-0)	0 (0-0)
			Megasphaera sp. Type 2	0 (0-0)	0 (0-0)
Mycoplasma hominis	2.8 (UD-6.5)	2.0 (UD-3.2)	*Mycoplasma hominis*	0.03 (0-0.3)	0 (0-0)
Parvimonas sp. Type 1	UD (UD-3.0)	UD (UD-UD)	*Parvimonas* spp.	0 (0-0.7)	0 (0-0)
Parvimonas sp. Type 2	UD (UD-7.0)	UD (UD-UD)			
Prevotella spp.	7.3 (4.8-8.8)	5.6 (4.2-7.4)	*Prevotella* spp.	0.5 (0-9.7)	0.05 (0-0.8)

Abbreviations: IQR, interquartile range; NGS, next generation sequencing; qPCR, quantitative polymerase chain reaction; sp., species (singular); spp., species (plural).
Data presented as median (IQR).
[a] Combined assay.

Figure 4.1. Hierarchically clustered Spearman correlations between (A) Concentrations of vaginal bacteria and (B) Relative abundances of vaginal bacteria in all women (N=110).



Abbreviations: sp., species (singular); spp., species (plural).

[a] Combined assay.

Table 4.3. Agreement in detection by qPCR and NGS in all women (N=110).

qPCR	n (%) detectable	NGS	n (%) detectable	Agreement	Expected agreement	Cohen's κ
Eggerthella sp. Type 1	51 (46%)	*Eggerthella* sp. Type 1	38 (35%)	86.4%	51.1%	0.72
Gemella asaccharolytica	52 (47%)	*Gemella asaccharolytica*	40 (36%)	87.3%	50.7%	0.74
Lactobacillus crispatus	26 (24%)	*Lactobacillus crispatus*	24 (22%)	94.6%	64.9%	0.84
Leptotrichia/Sneathia spp.	71 (65%)	*Leptotrichia/Sneathia* spp.	42 (38%)	71.8%	47.1%	0.47
Megasphaera spp. Types 1 & 2[a]	34 (31%)	*Megasphaera* sp. Type 1	13 (12%)	80.9%	64.6%	0.46
Megasphaera spp. Types 1 & 2[a]	34 (31%)	*Megasphaera* sp. Type 2	19 (17%)	84.6%	62.5%	0.59
Mycoplasma hominis	44 (40%)	*Mycoplasma hominis*	39 (35%)	95.5%	52.9%	0.90
Parvimonas sp. Type 1	28 (25%)	*Parvimonas* spp.	36 (33%)	76.4%	58.5%	0.43
Parvimonas sp. Type 2	31 (28%)	*Parvimonas* spp.	36 (33%)	90.0%	57.5%	0.76
Prevotella spp.	110 (100%)	*Prevotella* spp.	72 (65%)	65.5%	65.5%	--[b]

Abbreviations: NGS, next generation sequencing; qPCR, quantitative polymerase chain reaction; sp., species (singular); spp., species (plural).

[a] Combined assay.

[b] All samples were positive by qPCR for *Prevotella* spp.

Cohen's κ	Level of agreement
≤0	Poor
0-0.20	Slight
>0.20-0.40	Fair
>0.40-0.60	Moderate
>0.60-0.80	Substantial
>0.80	Almost perfect

Table 4.4. AUC for HIV acquisition by taxa concentration and relative abundance.

qPCR (concentration)	**AUC**	**95% CI**	**NGS (relative abundance)**	**AUC**	**95% CI**	**p-value**[b]
***Eggerthella* sp. Type 1**	**0.647**	**0.544-0.751**	***Eggerthella* sp. Type 1**	**0.647**	**0.543, 0.751**	0.993
Gemella asaccharolytica	**0.699**	**0.599, 0.798**	***Gemella asaccharolytica***	**0.719**	**0.622, 0.816**	0.472
Lactobacillus crispatus	0.526	0.417, 0.635	*Lactobacillus crispatus*	0.567	0.459, 0.675	0.408
***Leptotrichia/Sneathia* spp.**	**0.639**	**0.534, 0.743**	***Leptotrichia/Sneathia* spp.**	**0.645**	**0.542, 0.748**	0.858
***Megasphaera* spp. Types 1 & 2**[a]	**0.665**	**0.562, 0.767**	*Megasphaera* sp. Type 1	0.596	0.490, 0.703	0.205
	--	--	*Megasphaera* sp. Type 2	0.562	0.453, 0.670	0.051
Mycoplasma hominis	**0.646**	**0.542, 0.751**	***Mycoplasma hominis***	**0.635**	**0.529, 0.740**	0.691
***Parvimonas* sp. Type 1**	**0.623**	**0.517, 0.728**	***Parvimonas* spp.**	**0.656**	**0.553, 0.758**	0.510
***Parvimonas* sp. Type 2**	**0.667**	**0.565, 0.769**	***Parvimonas* spp.**	--	--	0.766
***Prevotella* spp.**	**0.636**	**0.532, 0.741**	***Prevotella* spp.**	**0.661**	**0.558, 0.764**	0.576
			Shannon's diversity index	**0.623**	**0.517, 0.728**	--

Abbreviations: AUC, area under the curve; CI, confidence interval; NGS, next generation sequencing; qPCR, quantitative polymerase chain reaction; sp., species (singular); spp., species (plural).

Bolded text indicates prediction values greater than chance alone.

[a] Combined assay.

[b] χ^2 test.

Table 4.5. HIV discrimination value from models using the first three principal components from principal components analysis (PCA).

PCA models	**AUC**	**95% CI**
(1) Demographics/Risk factors model (ref)[b]	0.706	0.609, 0.804
(2) qPCR (concentration) PCs alone[c,d]	0.709	0.609, 0.810
(3) NGS (relative abundance) PCs alone[e,f]	0.657	0.554, 0.760
(4) qPCR (concentration) PCs + ref model vars[g]	0.787	0.702, 0.872
(5) NGS (relative abundance) PCs + ref model vars[h]	0.761	0.672, 0.850

Comparison of PCA models		**AUCs**	**p-value**[a]
qPCR PCs alone vs. Demographics/Risk factors	(2 vs. 1)	0.709 vs. 0.706	0.964
NGS PCs alone vs. Demographics/Risk factors	(3 vs. 1)	0.657 vs. 0.706	0.498
qPCR PCs alone vs. NGS PCs alone	(2 vs. 3)	0.709 vs. 0.657	0.226
qPCR PCs + ref model vs. Demographics/Risk factors	(4 vs. 1)	0.787 vs. 0.706	0.076
NGS PCs + ref model vs. Demographics/Risk factors	(5 vs. 1)	0.761 vs. 0.706	0.173
qPCR PCs + ref model vs. NGS PCs + ref model	(4 vs. 5)	0.787 vs. 0.761	0.372

Abbreviations: AUC, area under the curve; CI, confidence interval; NGS, next generation sequencing; PC, principal components; PCA, principal components analysis; qPCR, quantitative polymerase chain reaction.

[a] χ^2 test.

[b] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, and recent condomless sex.

[c] Kaiser–Meyer–Olkin value=0.9136; first 3 PCs accounted for 83.92% of the variance in the qPCR data.

[d] Adjusted for matching factors and first 3 qPCR PCs.

[e] Kaiser–Meyer–Olkin value=0.7399; first 3 PCs accounted for 63.13% of the variance in the NGS data.

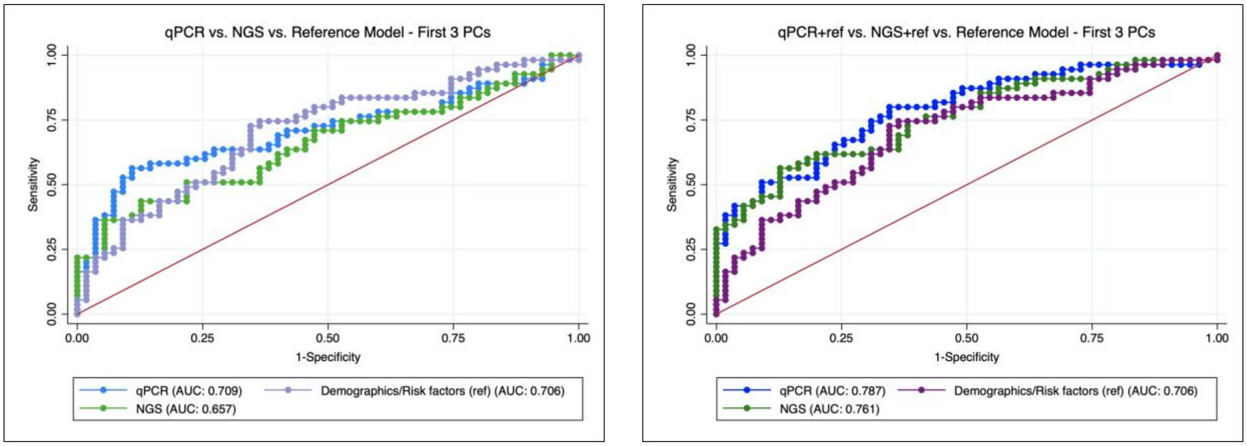
[f] Adjusted for matching factors and first 3 NGS PCs.

[g] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, recent condomless sex, and first 3 qPCR PCs.

[h] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, recent condomless sex, and first 3 NGS PCs.

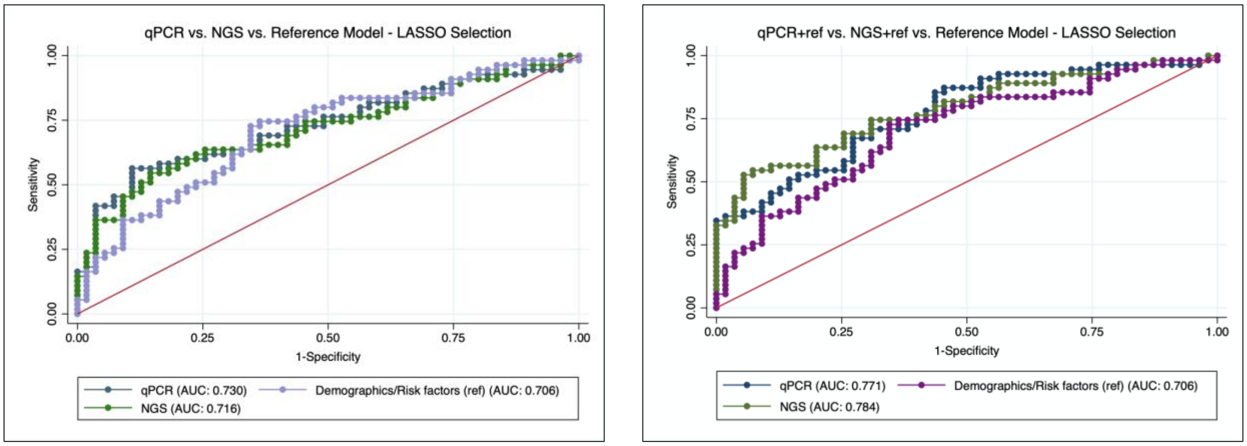
Figure 4.2. Receiver-operating curves for HIV discrimination from: (A,B) models using the first three principal components from principal components analysis (PCA) and (C,D) models selected using least absolute shrinkage and selection operator (LASSO).

PCA


A. PCR data alone

B. PCR data + Reference model

LASSO


C. PCR data alone

D. PCR data with Reference model

Abbreviations: AUC, area under the curve; LASSO, least absolute shrinkage and selection operator; NGS, next generation sequencing; PCA, principal components analysis; PCR, polymerase chain reaction; qPCR, quantitative polymerase chain reaction.

Table 4.6. HIV discrimination value from models selected using least absolute shrinkage and selection operator (LASSO).

LASSO models	**AUC**	**95% CI**
(1) Demographics/Risk factors model (ref)[b]	0.706	0.609, 0.804
(2) qPCR (concentration) selected vars[c,d]	0.730	0.634, 0.826
(3) NGS (relative abundance) selected vars[e,f]	0.716	0.619, 0.813
(4) qPCR (concentration) selected vars (includes ref vars)[g,h]	0.771	0.684, 0.858
(5) NGS (relative abundance) selected vars (includes ref vars)[i,j]	0.784	0.698, 0.870

Comparison of LASSO models		**AUCs**	**p-value[a]**
qPCR selected vars alone vs. Demographics/Risk factors	(2 vs. 1)	0.730 vs. 0.706	0.737
NGS selected vars alone vs. Demographics/Risk factors	(3 vs. 1)	0.716 vs. 0.706	0.883
qPCR selected vars alone vs. NGS selected vars alone	(2 vs. 3)	0.730 vs. 0.716	0.758
qPCR selected vars (inc. ref vars) vs. Demographics/Risk factors	(4 vs. 1)	0.771 vs. 0.706	0.213
NGS selected vars (inc. ref vars) vs. Demographics/Risk factors	(5 vs. 1)	0.784 vs. 0.706	0.151
qPCR selected vars (inc. ref vars) vs. NGS selected vars (inc. ref vars)	(4 vs. 5)	0.771 vs. 0.784	0.726

Abbreviations: AUC, area under the curve; CI, confidence interval; LASSO, least absolute shrinkage and selection operator; NGS, next generation sequencing; qPCR, quantitative polymerase chain reaction.

[a] χ^2 test.

[b] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, and recent condomless sex.

[c] Selected variables: *G. asaccharolytica*, *L. crispatus*, *Megasphaera* spp. Types 1 & 2, *M. hominis*, *Parvimonas* sp. Type 1, and *Parvimonas* sp. Type 2.

[d] Adjusted for matching factors and selected variables in (C).

[e] Selected variables: *Lactobacillus iners*, *Leptotrichia amnionii*, *Gemella* spp., *Eggerthella* sp. Type 1, and *Escherichia coli*.

[f] Adjusted for matching factors and selected variables in (e).

[g] Selected variables: *G. asaccharolytica*, *L. crispatus*, *Megasphaera* spp. Types 1 & 2, *Parvimonas* sp. Type 1, *Parvimonas* sp. Type 2, age, and pregnancy/contraceptive status.

[h] Adjusted for matching factors and selected variables in (g).

[i] Selected variables: *L. iners*, *L. amnionii*, *Gemella* spp., *Eggerthella* sp. Type 1, *E. coli*, and pregnancy/contraceptive status.

[j] Adjusted for matching factors and selected variables in (i).

Table S4.1. Agreement for HIV acquisition between models – principal components analysis (PCA).

	n (%) predicted positive		n (%) predicted positive	Agreement	Expected agreement	Cohen's κ
Demographics/Risk factors model (ref)[a]	53 (48.2%)	qPCR PCs alone[b]	51 (46.4%)	45.5%	50.1%	-0.0938 (Poor)
Demographics/Risk factors model (ref)[a]	53 (48.2%)	NGS PCs alone[c]	40 (36.4%)	50.0%	50.5%	-0.0100 (Poor)
qPCR PCs alone[b]	51 (46.4%)	NGS PCs alone[c]	40 (36.4%)	82.7%	51.0%	0.6476 (Substantial)
Demographics/Risk factors model (ref)[a]	53 (48.2%)	qPCR PCs + ref model vars[d]	52 (47.3%)	66.4%	50.1%	0.3259 (Fair)
Demographics/Risk factors model (ref)[a]	53 (48.2%)	NGS PCs + ref model vars[e]	54 (49.1%)	71.8%	50.0%	0.4360 (Moderate)
qPCR PCs + ref model vars[d]	52 (47.3%)	NGS PCs + ref model vars[e]	54 (49.1%)	81.8%	51.0%	0.6360 (Substantial)

Abbreviations: NGS, next generation sequencing; PC, principal components; PCA, principal components analysis; qPCR, quantitative polymerase chain reaction.

[a] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, and recent condomless sex.

[b] Adjusted for matching factors and first 3 qPCR PCs.

[c] Adjusted for matching factors and first 3 NGS PCs.

[d] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, recent condomless sex, and first 3 qPCR PCs.

[e] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, recent condomless sex, and first 3 NGS PCs.

Table S4.2. Agreement for HIV acquisition between models – least absolute shrinkage and selection operator (LASSO).

	n (%) predicted positive		n (%) predicted positive	Agreement	Expected agreement	Cohen's κ
Demographics/Risk factors model (ref)[a]	53 (48.2%)	qPCR selected vars[b,c]	45 (40.9%)	52.7%	50.3%	0.0483 (Slight)
Demographics/Risk factors model (ref)[a]	53 (48.2%)	NGS selected vars[d,e]	51 (46.4%)	52.7%	50.1%	0.0520 (Slight)
qPCR (concentration) selected vars[b,c]	45 (40.9%)	NGS selected vars[d,e]	51 (46.4%)	81.8%	50.7%	0.6315 (Substantial)
Demographics/Risk factors model (ref)[a]	53 (48.2%)	qPCR selected vars (inc. ref vars)[f,g]	47 (42.7%)	58.2%	50.3%	0.1592 (Slight)
Demographics/Risk factors model (ref)[a]	53 (48.2%)	NGS selected vars (inc. ref vars)[h,i]	44 (40.0%)	64.6%	50.4%	0.2857 (Fair)
qPCR selected vars (inc. ref vars)[f,g]	47 (42.7%)	NGS selected vars (inc. ref vars)[h,i]	44 (40.0%)	75.5%	51.5%	0.4944 (Moderate)

Abbreviations: NGS, next generation sequencing; qPCR, quantitative polymerase chain reaction.

[a] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, and recent condomless sex.

[b] Selected variables: *G. asaccharolytica*, *L. crispatus*, *Megasphaera* spp. Types 1 & 2, *M. hominis*, *Parvimonas* sp. Type 1, and *Parvimonas* sp. Type 2.

[c] Adjusted for matching factors and selected variables in (b).

[d] Selected variables: *Lactobacillus iners*, *Leptotrichia amnionii*, *Gemella* spp., *Eggerthella* sp. Type 1, and *Escherichia coli*.

[e] Adjusted for matching factors and selected variables in (d).

[f] Selected variables: *G. asaccharolytica*, *L. crispatus*, *Megasphaera* spp. Types 1 & 2, *Parvimonas* sp. Type 1, *Parvimonas* sp. Type 2, age, and pregnancy/contraceptive status.

[g] Adjusted for matching factors and selected variables in (f).

[h] Selected variables: *L. iners*, *L. amnionii*, *Gemella* spp., *Eggerthella* sp. Type 1, *E. coli*, and pregnancy/contraceptive status.

[i] Adjusted for matching factors and selected variables in (h).

Table S4.3. Detection of bacteria associated with HIV acquisition via qPCR and NGS in all women (N=110).

qPCR			**NGS**		
	Cases (n=55)	**Controls (n=55)**		**Cases (n=55)**	**Controls (n=55)**
Eggerthella sp. Type 1	32 (58%)	19 (35%)	*Eggerthella* sp. Type 1	24 (44%)	14 (25%)
Gemella asaccharolytica	33 (60%)	19 (35%)	*Gemella asaccharolytica*	30 (55%)	10 (18%)
Lactobacillus crispatus	15 (27%)	11 (20%)	*Lactobacillus crispatus*	13 (24%)	11 (20%)
Leptotrichia/Sneathia spp.	39 (71%)	32 (58%)	*Leptotrichia/Sneathia* spp.	29 (53%)	15 (27%)
Megasphaera spp. Types 1 & 2[a]	25 (45%)	9 (16%)	*Megasphaera* sp. type 1	11 (20%)	2 (4%)
			Megasphaera sp. type 2	12 (22%)	7 (13%)
Mycoplasma hominis	28 (51%)	16 (29%)	*Mycoplasma hominis*	28 (51%)	11 (20%)
Parvimonas sp. Type 1	20 (36%)	8 (15%)	*Parvimonas* spp.	26 (47%)	10 (18%)
Parvimonas sp. Type 2	24 (44%)	7 (13%)			
Prevotella spp.	55 (100%)	55 (100%)	*Prevotella* spp.	35 (64%)	37 (67%)

Abbreviations: NGS, next generation sequencing; qPCR, quantitative polymerase chain reaction; sp., species (singular); spp., species (plural).
Data presented as n (%).
[a] Combined assay.

Table S4.4. AUC for HIV acquisition by taxa detection via qPCR and NGS.

qPCR	**AUC**	**95% CI**	**NGS**	**AUC**	**95% CI**	**p-value[b]**
Eggerthella sp. Type 1	**0.636**	**0.532, 0.740**	**_Eggerthella_ sp. Type 1**	**0.614**	**0.509, 0.720**	0.590
Gemella asaccharolytica	**0.661**	**0.558, 0.764**	**_Gemella asaccharolytica_**	**0.715**	**0.617, 0.812**	0.167
Lactobacillus crispatus	0.548	0.439, 0.656	*Lactobacillus crispatus*	0.533	0.423, 0.642	0.596
Leptotrichia/Sneathia spp.	0.581	0.473, 0.689	**_Leptotrichia/Sneathia_ spp.**	**0.653**	**0.550, 0.756**	0.158
Megasphaera spp. Types 1 & 2[a]	**0.672**	**0.570, 0.773**	*Megasphaera* sp. type 1	**0.610**	**0.504, 0.716**	0.260
	--	--	*Megasphaera* sp. type 2	0.578	0.471, 0.685	0.049
Mycoplasma hominis	**0.621**	**0.515, 0.726**	**_Mycoplasma hominis_**	**0.662**	**0.559, 0.765**	0.156
Parvimonas sp. Type 1	**0.633**	**0.528, 0.738**	**_Parvimonas_ spp.**	**0.669**	**0.568, 0.770**	0.498
Parvimonas sp. Type 2	**0.677**	**0.576, 0.777**	**_Parvimonas_ spp.**	--	--	0.850
Prevotella spp.	0.524	0.415, 0.634	*Prevotella* spp.	0.538	0.429, 0.646	0.754
			Shannon's diversity index	**0.623**	**0.517, 0.728**	--

Abbreviations: AUC, area under the curve; CI, confidence interval; NGS, next generation sequencing; qPCR, quantitative polymerase chain reaction; sp., species (singular); spp., species (plural).

Bolded text indicates prediction values greater than chance alone.

[a] Combined assay.

[b] χ^2 test.

Table S4.5. HIV discrimination value from detection models using the first three principal components from principal components analysis (PCA).

PCA models		**AUC**	**95% CI**
(1) Demographics/Risk factors model (ref)[b]		0.706	0.609, 0.804
(2) qPCR (detectable) PCs alone[c,d]		0.710	0.612, 0.808
(3) NGS (detectable) PCs alone[e,f]		0.707	0.609, 0.806
(4) qPCR (detectable) PCs + ref model vars[g]		0.775	0.688, 0.862
(5) NGS (detectable) PCs + ref model vars[h]		0.767	0.679, 0.855
Comparison of PCA models		**AUCs**	**p-value[a]**
qPCR PCs alone vs. Demographics/Risk factors	(2 vs. 1)	0.710 vs. 0.706	0.958
NGS PCs alone vs. Demographics/Risk factors	(3 vs. 1)	0.707 vs. 0.706	0.985
qPCR PCs alone vs. NGS PCs alone	(2 vs. 3)	0.710 vs. 0.707	0.947
qPCR PCs + ref model vs. Demographics/Risk factors	(4 vs. 1)	0.775 vs. 0.706	0.106
NGS PCs + ref model vs. Demographics/Risk factors	(5 vs. 1)	0.767 vs. 0.706	0.141
qPCR PCs + ref model vs. NGS PCs + ref model	(4 vs. 5)	0.775 vs. 0.767	0.765

Abbreviations: AUC, area under the curve; CI, confidence interval; NGS, next generation sequencing; PC, principal components; PCA, principal components analysis; qPCR, quantitative polymerase chain reaction.

[a] χ^2 test.

[b] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, and recent condomless sex.

[c] Kaiser–Meyer–Olkin value=0.8435; first 3 PCs accounted for 73.92% of the variance in the qPCR data.

[d] Adjusted for matching factors and first 3 qPCR PCs.

[e] Kaiser–Meyer–Olkin value=0.9110; first 3 PCs accounted for 73.90% of the variance in the NGS data.

[f] Adjusted for matching factors and first 3 NGS PCs.

[g] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, recent condomless sex, and first 3 qPCR PCs.

[h] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, recent condomless sex, and first 3 NGS PCs.

Table S4.6. HIV discrimination value from detection models selected using least absolute shrinkage and selection operator (LASSO).

LASSO models	**AUC**	**95% CI**
(1) Demographics/Risk factors model (ref)[b]	0.706	0.609, 0.804
(2) qPCR (detectable) selected vars[c,d]	0.720	0.623, 0.816
(3) NGS (detectable) selected vars[e,f]	0.797	0.714, 0.880
(4) qPCR (detectable) selected vars (including ref vars)[g,h]	0.761	0.672, 0.850
(5) NGS (detectable) selected vars (including ref vars)[i,j]	0.801	0.719, 0.882

Comparison of LASSO models		**AUCs**	**p-value[a]**
qPCR selected vars alone vs. Demographics/Risk factors	(2 vs. 1)	0.720 vs. 0.706	0.846
NGS selected vars alone vs. Demographics/Risk factors	(3 vs. 1)	0.797 vs. 0.706	0.162
qPCR selected vars alone vs. NGS selected vars alone	(2 vs. 3)	0.720 vs. 0.797	0.103
qPCR selected vars (inc. ref vars) vs. Demographics/Risk factors	(4 vs. 1)	0.761 vs. 0.706	0.280
NGS selected vars (inc. ref vars) vs. Demographics/Risk factors	(5 vs. 1)	0.801 vs. 0.706	0.071
qPCR selected vars (inc. ref vars) vs. NGS selected vars (inc. ref vars)	(4 vs. 5)	0.761 vs. 0.801	0.291

Abbreviations: AUC, area under the curve; CI, confidence interval; LASSO, least absolute shrinkage and selection operator; NGS, next generation sequencing; qPCR, quantitative polymerase chain reaction.

[a] χ^2 test.

[b] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, and recent condomless sex.

[c] Selected variables: *L. crispatus*, *Megasphaera* spp. Types 1 & 2, *Parvimonas* sp. Type 1, and *Parvimonas* sp. Type 2.

[d] Adjusted for matching factors and selected variables in (C).

[e] Selected variables: *Gemella* spp., *L. crispatus*, *Megasphaera* sp. Type 1, *M. hominis*, *Parvimonas micra*, and *Prevotella* spp.

[f] Adjusted for matching factors and selected variables in (e).

[g] Selected variables: *L. crispatus*, *Megasphaera* spp. Types 1 & 2, *Parvimonas* sp. Type 1, *Parvimonas* sp. Type 2, and pregnancy/contraceptive status.

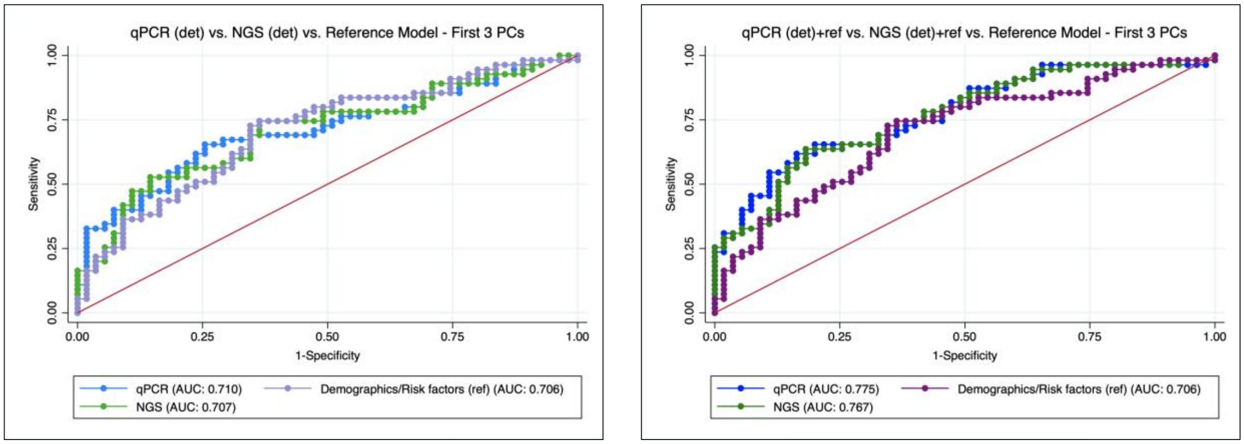
[h] Adjusted for matching factors and selected variables in (g).

[i] Selected variables: *Gemella* spp., *L. crispatus*, *Megasphaera* sp. Type 1, *M. hominis*, *Parvimonas micra*, and *Prevotella* spp., age, and pregnancy/contraceptive status.

[j] Adjusted for matching factors and selected variables in (i).

Figure S4.1. Receiver-operating curves for HIV discrimination from: (A,B) detection models using the first three principal components from principal components analysis (PCA) and (C,D) detection models selected using least absolute shrinkage and selection operator (LASSO).

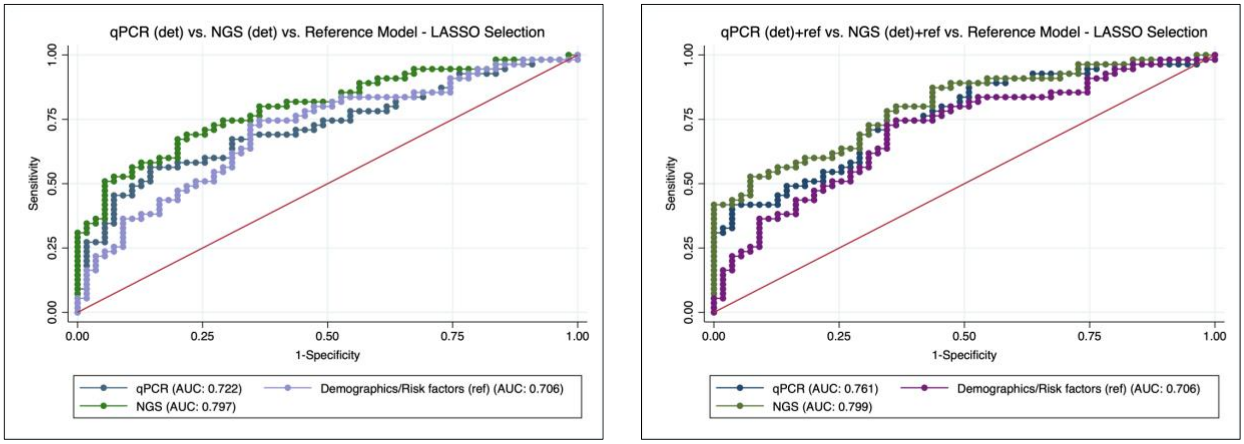
PCA



A. PCR data alone

B. PCR data + Reference model

LASSO



C. PCR data alone

D. PCR data with Reference model

Abbreviations: AUC, area under the curve; LASSO, least absolute shrinkage and selection operator; NGS, next generation sequencing; PCA, principal components analysis; PCR, polymerase chain reaction; qPCR, quantitative polymerase chain reaction.

Table S4.7. Agreement for HIV acquisition between detection models – principal components analysis (PCA).

	n (%) predicted positive		n (%) predicted positive	Agreement	Expected agreement	Cohen's κ
Demographics/Risk factors model (ref)[a]	53 (48.2%)	qPCR PCs alone[b]	51 (46.4%)	50.9%	50.1%	0.0156 (Slight)
Demographics/Risk factors model (ref)[a]	53 (48.2%)	NGS PCs alone[c]	55 (50.0%)	50.9%	50.0%	0.0182 (Slight)
qPCR PCs alone[b]	51 (46.4%)	NGS PCs alone[c]	55 (50.0%)	81.8%	50.0%	0.6364 (Substantial)
Demographics/Risk factors model (ref)[a]	53 (48.2%)	qPCR PCs + ref model vars[d]	48 (43.6%)	73.6%	50.2%	0.4703 (Moderate)
Demographics/Risk factors model (ref)[a]	53 (48.2%)	NGS PCs + ref model vars[e]	50 (45.5%)	75.5%	50.2%	0.5075 (Moderate)
qPCR PCs + ref model vars[d]	48 (43.6%)	NGS PCs + ref model vars[e]	50 (45.5%)	85.5%	50.6%	0.7057 (Substantial)

Abbreviations: NGS, next generation sequencing; PC, principal components; PCA, principal components analysis; qPCR, quantitative polymerase chain reaction.

[a] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, and recent condomless sex.

[b] Adjusted for matching factors and first 3 qPCR PCs.

[c] Adjusted for matching factors and first 3 NGS PCs.

[d] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, recent condomless sex, and first 3 qPCR PCs.

[e] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, recent condomless sex, and first 3 NGS PCs.

Table S4.8. Agreement for HIV acquisition between detection models – least absolute shrinkage and selection operator (LASSO).

	n (%) predicted positive		**n (%) predicted positive**	**Agreement**	**Expected agreement**	**Cohen's κ**
Demographics/Risk factors model (ref)[a]	53 (48.2%)	qPCR selected vars[b,c]	45 (40.9%)	52.7%	50.3%	0.0483 (Slight)
Demographics/Risk factors model (ref)[a]	53 (48.2%)	NGS selected vars[d,e]	49 (44.6%)	56.4%	50.2%	0.1238 (Slight)
qPCR (concentration) selected vars[b,c]	45 (40.9%)	NGS selected vars[d,e]	49 (44.6%)	74.6%	51.0%	0.4806 (Moderate)
Demographics/Risk factors model (ref)[a]	53 (48.2%)	qPCR selected vars (inc. ref vars)[f,g]	45 (40.9%)	65.5%	50.3%	0.3045 (Fair)
Demographics/Risk factors model (ref)[a]	53 (48.2%)	NGS selected vars (inc. ref vars)[h,i]	54 (49.1%)	68.2%	50.0%	0.3632 (Fair)
qPCR selected vars (inc. ref vars)[f,g]	45 (40.9%)	NGS selected vars (inc. ref vars)[h,i]	54 (49.1%)	73.6%	50.2%	0.4710 (Moderate)

Abbreviations: NGS, next generation sequencing; qPCR, quantitative polymerase chain reaction.

[a] Adjusted for matching factors, age, pregnancy/contraceptive status, number of recent sexual partners, frequency of sex in the past month, and recent condomless sex.

[b] Selected variables: *L. crispatus*, *Megasphaera* spp. Types 1 & 2, *Parvimonas* sp. Type 1, and *Parvimonas* sp. Type 2.

[c] Adjusted for matching factors and selected variables in (b).

[d] Selected variables: *Gemella* spp., *L. crispatus*, *Megasphaera* sp. Type 1, *M. hominis*, *Parvimonas micra*, and *Prevotella* spp.

[e] Adjusted for matching factors and selected variables in (d).

[f] Selected variables: *L. crispatus*, *Megasphaera* spp. Types 1 & 2, *Parvimonas* sp. Type 1, *Parvimonas* sp. Type 2, and pregnancy/contraceptive status.

[g] Adjusted for matching factors and selected variables in (f).

[h] Selected variables: *Gemella* spp., *L. crispatus*, *Megasphaera* sp. Type 1, *M. hominis*, *Parvimonas micra*, and *Prevotella* spp., age, and pregnancy/contraceptive status.

[i] Adjusted for matching factors and selected variables in (h).

CHAPTER 5. Conclusion

In conclusion, these findings suggest that nonoptimal vaginal bacterial communities are common among Kenyan women, and that DMPA-IM use among postpartum women may be associated with maintenance of highly diverse, nonoptimal vaginal bacterial communities. Additionally, specific bacteria associated with increased HIV acquisition risk were present among women at average risk for HIV, and that DMPA-IM use, compared to non-HC user, showed differential patterns in concentration change with three taxa associated with HIV acquisition, especially among a subset of women with these species already present.

While the PFP cohort was small and restricted to a specific population of women, these findings suggest that DMPA-IM's effect on the vaginal microenvironment deserves further attention, especially in postpartum women. DMPA-IM is suitable for use in breastfeeding women four or more weeks postpartum because progestins do not suppress production of breast milk,[7] making it one of the few approved highly effective contraceptives in the postpartum period. Additionally, while the ECHO trial has relieved apprehension about use of DMPA-IM in women in high HIV prevalence settings,[31] nonoptimal vaginal microbiota are a risk factor for other negative health outcomes, including PID and the acquisition of other STIs; DMPA-IM use in the postpartum period may pose greater risks than it does in the non-postpartum period, due to high vaginal microbial diversity seen in this period and the potential for DMPA-IM to stabilize nonoptimal vaginal composition.

We found that both the absolute concentrations and relative abundances of bacterial taxa associated with HIV acquisition perform equally well as predictors for HIV acquisition; the approaches may be similar in utility for epidemiologic investigations relating the sexually transmitted infection risk. For the purposes of understanding the microbiome and for discovery, NGS has a distinct advantage; however, it is more labor intensive and requires specialized laboratory facilities. Alternatively, qPCR is easier to deploy to non-research lab settings and may be more broadly useful in real-world scenarios. Depending on setting and purpose one method may have benefit over the other, but for measuring risk of HIV acquisition, both methods appear to be useful tools.

Acknowledgements

I would like to thank my dissertation chair, Brandon Guthrie, for his unwavering support, guidance. I thank Alison Roxby, my mentor and role model, for her encouragement, guidance, and for providing me with numerous professional opportunities in the field of HIV/STI research. I would also like to thank the other members of my dissertation committee, R. Scott McClelland and Jim Hughes, for their time and invaluable guidance and feedback. I could not have finished this endeavor without the continuous support, cheerleading, and guidance from Joseph Chris Delaney and Heidi Crane; they provided me invaluable opportunities, endless training, and a wonderful working environment.

Numerous others supported this work and deserve recognition. These studies would not be possible without the the women who participated in the study as well as the staff at the Mathare North Health Center in Nairobi, Kenya. Many others contributed to work presented in this dissertation, including Kenneth Tapia, Sujatha Srinivasan, David N Fredricks, Eric Munene Muriuki, Bhavna Chohan, Jacqueline Wallis, Congzhou Liu, Jairam Lingappa, and Grace John-Stewart.

I would also like to acknowledge the funding sources that supported this work. The research in Chapters 2 and 3 was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health [R21AI129712; PI: Alison Roxby], the National Institute of Child Health and Human Development of the National Institutes of Health [K23HD071788; PI: Alison Roxby], and a University of Washington Royalty Research Fund [A106982; PI: Alison Roxby]. The research in Chapter 4 was supported by the National Institute of Child Health and Human Development of the National Institutes of Health [P01HD64915;

project lead: McClelland]. I was partially supported by the STD/AIDS Research Training Fellowship Program (2018-2019) (NIH T32AI007140; PI: Sheila Lukehart).

Finally I would like to thank my PhD cohort members, friends, and most importantly my family, for their unwavering support.

References

1. WHO. Family planning/Contraception. 2018; http://www.who.int/mediacentre/factsheets/fs351/en/.

2. *Adding It Up: Investing in Contraception and Maternal and Newborn Health, 2017*: Guttmacher Institute;2017.

3. van de Wijgert J, Jespers V. The global health impact of vaginal dysbiosis. *Res Microbiol.* Nov - Dec 2017;168(9-10):859-864.

4. *UNAIDS data 2018.* UNAIDS: UNAIDS;2018.

5. Ross JA, Agwanda AT. Increased use of injectable contraception in sub-Saharan Africa. *Afr J Reprod Health.* Dec 2012;16(4):68-80.

6. American College of Obstetricians and Gynecologists CoAHC, Committee on Gynecologic Practice. Depot medroxyprogesterone acetate and bone effects. Committee Opinion No. 602. *Obstet Gynecol.* 2014(123):1398–1402.

7. Curtis KM, Tepper NK, Jatlaoui TC, et al. U.S. Medical Eligibility Criteria for Contraceptive Use, 2016. *MMWR Recomm Rep.* Jul 29 2016;65(3):1-103.

8. *National Family Planning Guidelines for Service Providers 6th Edition*: Kenyan Ministry of Health;2018.

9. DEPO-PROVERA - medroxyprogesterone acetate injection, suspension 2006; https://dailymed.nlm.nih.gov/dailymed/archives/fdaDrugInfo.cfm?archiveid=11565.

10. Baeten JM, Heffron R. Contraception and sexually transmitted infections: risks and benefits, hypotheses and evidence. *Lancet Glob Health.* Aug 2015;3(8):e430-e431.

11. Ravel J, Gajer P, Abdo Z, et al. Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A.* Mar 15 2011;108 Suppl 1:4680-4687.

12. Smith SB, Ravel J. The vaginal microbiota, host defence and reproductive physiology. *J Physiol.* Jan 15 2017;595(2):451-463.

13. McKinnon LR, Achilles SL, Bradshaw CS, et al. The Evolving Facets of Bacterial Vaginosis: Implications for HIV Transmission. *AIDS Res Hum Retroviruses.* Mar 2019;35(3):219-228.

14. Anahtar MN, Byrne EH, Doherty KE, et al. Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. *Immunity.* May 19 2015;42(5):965-976.

15. Brotman RM. Vaginal microbiome and sexually transmitted infections: an epidemiologic perspective. *J Clin Invest.* Dec 2011;121(12):4610-4617.

16. Gosmann C, Anahtar MN, Handley SA, et al. Lactobacillus-Deficient Cervicovaginal Bacterial Communities Are Associated with Increased HIV Acquisition in Young South African Women. *Immunity.* Jan 17 2017;46(1):29-37.

17. 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.* May 2018;18(5):554-564.

18. Atashili J, Poole C, Ndumbe PM, Adimora AA, Smith JS. Bacterial vaginosis and HIV acquisition: a meta-analysis of published studies. *AIDS.* Jul 31 2008;22(12):1493-1501.

19. Balkus JE, Manhart LE, Lee J, et al. Periodic Presumptive Treatment for Vaginal Infections May Reduce the Incidence of Sexually Transmitted Bacterial Infections. *J Infect Dis.* Jun 15 2016;213(12):1932-1937.

20. Schwebke JR, Desmond R. A randomized trial of metronidazole in asymptomatic bacterial vaginosis to prevent the acquisition of sexually transmitted diseases. *Am J Obstet Gynecol.* Jun 2007;196(6):517 e511-516.

21. van de Wijgert JH, Borgdorff H, Verhelst R, et al. The vaginal microbiota: what have we learned after a decade of molecular characterization? *PLoS One.* 2014;9(8):e105998.

22. Brotman RM, Ravel J, Bavoil PM, Gravitt PE, Ghanem KG. Microbiome, sex hormones, and immune responses in the reproductive tract: challenges for vaccine development against sexually transmitted infections. *Vaccine.* Mar 20 2014;32(14):1543-1552.

23. Hapgood JP, Kaushic C, Hel Z. Hormonal Contraception and HIV-1 Acquisition: Biological Mechanisms. *Endocr Rev.* Feb 1 2018;39(1):36-78.

24. 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.

25. Miller L, Patton DL, Meier A, Thwin SS, Hooton TM, Eschenbach DA. Depomedroxyprogesterone-induced hypoestrogenism and changes in vaginal flora and epithelium. *Obstet Gynecol.* Sep 2000;96(3):431-439.

26. Mirmonsef P, Hotton AL, Gilbert D, et al. Free glycogen in vaginal fluids is associated with Lactobacillus colonization and low vaginal pH. *PLoS One.* 2014;9(7):e102467.

27. Mirmonsef P, Modur S, Burgad D, et al. Exploratory comparison of vaginal glycogen and Lactobacillus levels in premenopausal and postmenopausal women. *Menopause.* Jul 2015;22(7):702-709.

28. Gajer P, Brotman RM, Bai G, et al. Temporal dynamics of the human vaginal microbiota. *Sci Transl Med.* May 2 2012;4(132):132ra152.

29. Polis CB, Curtis KM, Hannaford PC, et al. An updated systematic review of epidemiological evidence on hormonal contraceptive methods and HIV acquisition in women. *AIDS.* Nov 13 2016;30(17):2665-2683.

30. Murphy K, Irvin SC, Herold BC. Research gaps in defining the biological link between HIV risk and hormonal contraception. *Am J Reprod Immunol.* Aug 2014;72(2):228-235.

31. Ahmed K, Baeten JM, Beksinska M, et al. HIV incidence among women using intramuscular depot medroxyprogesterone acetate, a copper intrauterine device, or a levonorgestrel implant for contraception: a randomised, multicentre, open-label trial. *The Lancet.* 2019.

32. Hapgood H. Is the injectable contraceptive Depo-medroxyprogesterone acetate (DMPA-IM) associated with an increased risk for HIV acquisition? The jury is still out. *AIDS Res Hum Retroviruses.* Dec 4 2019.

33. Baeten JM, Nyange PM, Richardson BA, et al. Hormonal contraception and risk of sexually transmitted disease acquisition: results from a prospective study. *Am J Obstet Gynecol.* Aug 2001;185(2):380-385.

34. McClelland RS, Richardson BA, Graham SM, et al. A prospective study of risk factors for bacterial vaginosis in HIV-1-seronegative African women. *Sex Transm Dis.* Jun 2008;35(6):617-623.

35. Pettifor A, Delany S, Kleinschmidt I, Miller WC, Atashili J, Rees H. Use of injectable progestin contraception and risk of STI among South African women. *Contraception.* Dec 2009;80(6):555-560.

36. van de Wijgert JH, Verwijs MC, Turner AN, Morrison CS. Hormonal contraception decreases bacterial vaginosis but oral contraception may increase candidiasis: implications for HIV transmission. *AIDS.* Aug 24 2013;27(13):2141-2153.

37. Brooks JP, Edwards DJ, Blithe DL, et al. Effects of combined oral contraceptives, depot medroxyprogesterone acetate and the levonorgestrel-releasing intrauterine system on the vaginal microbiome. *Contraception.* Apr 2017;95(4):405-413.

38. Achilles SL, Austin MN, Meyn LA, Mhlanga F, Chirenje ZM, Hillier SL. Impact of contraceptive initiation on vaginal microbiota. *Am J Obstet Gynecol.* Jun 2018;218(6):622 e621-622 e610.

39. Mitchell CM, McLemore L, Westerberg K, et al. Long-term effect of depot medroxyprogesterone acetate on vaginal microbiota, epithelial thickness and HIV target cells. *J Infect Dis.* Aug 15 2014;210(4):651-655.

40. Roxby AC, Fredricks DN, Odem-Davis K, et al. Changes in Vaginal Microbiota and Immune Mediators in HIV-1-Seronegative Kenyan Women Initiating Depot Medroxyprogesterone Acetate. *J Acquir Immune Defic Syndr.* Apr 1 2016;71(4):359-366.

41. Birse KD, Romas LM, Guthrie BL, et al. Genital Injury Signatures and Microbiome Alterations Associated With Depot Medroxyprogesterone Acetate Usage and Intravaginal Drying Practices. *J Infect Dis.* Feb 15 2017;215(4):590-598.

42. 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.* Sep 20 2017;7(1):11974.

43. Yang L, Hao Y, Hu J, et al. Differential effects of depot medroxyprogesterone acetate administration on vaginal microbiome in Hispanic White and Black women. *Emerg Microbes Infect.* 2019;8(1):197-210.

44. Byrne EH, Anahtar MN, Cohen KE, et al. Association between injectable progestin-only contraceptives and HIV acquisition and HIV target cell frequency in the female genital tract in South African women: a prospective cohort study. *The Lancet Infectious Diseases.* 2016;16(4):441-448.

45. Fredricks DN. Molecular methods to describe the spectrum and dynamics of the vaginal microbiota. *Anaerobe.* Aug 2011;17(4):191-195.

46. Golob JL, Margolis E, Hoffman NG, Fredricks DN. Evaluating the accuracy of amplicon-based microbiome computational pipelines on simulated human gut microbial communities. *BMC Bioinformatics.* May 30 2017;18(1):283.

47. Clooney AG, Fouhy F, Sleator RD, et al. Comparing Apples and Oranges?: Next Generation Sequencing and Its Impact on Microbiome Analysis. *PLoS One.* 2016;11(2):e0148028.

48. Lewis FM, Bernstein KT, Aral SO. Vaginal Microbiome and Its Relationship to Behavior, Sexual Health, and Sexually Transmitted Diseases. *Obstet Gynecol.* Apr 2017;129(4):643-654.

49. Mitchell C, Marrazzo J. Bacterial vaginosis and the cervicovaginal immune response. *Am J Reprod Immunol.* Jun 2014;71(6):555-563.

50. van de Wijgert J. The vaginal microbiome and sexually transmitted infections are interlinked: Consequences for treatment and prevention. *PLoS Med.* Dec 2017;14(12):e1002478.

51. Mishell DR, Jr. Pharmacokinetics of depot medroxyprogesterone acetate contraception. *J Reprod Med.* May 1996;41(5 Suppl):381-390.

52. Ortiz A, Hirol M, Stanczyk FZ, Goebelsmann U, Mishell DR. Serum medroxyprogesterone acetate (MPA) concentrations and ovarian function following intramuscular injection of depo-MPA. *J Clin Endocrinol Metab.* Jan 1977;44(1):32-38.

53. Wessels JM, Lajoie J, Cooper M, et al. Medroxyprogesterone acetate alters the vaginal microbiota and microenvironment in women and increases susceptibility to HIV-1 in humanized mice. *Dis Model Mech.* Oct 23 2019;12(10).

54. Pharmacokinetic Profiling (Biomarkers Core Laboratory), Columbia University Irving Medical Center. https://www.irvinginstitute.columbia.edu/services/pharmacokinetic-profiling.

55. Nugent RP, Krohn MA, Hillier SL. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *J Clin Microbiol.* Feb 1991;29(2):297-301.

56. Khot PD, Ko DL, Hackman RC, Fredricks DN. Development and optimization of quantitative PCR for the diagnosis of invasive aspergillosis with bronchoalveolar lavage fluid. *BMC Infect Dis.* 2008;8:73.

57. Srinivasan S, Hoffman NG, Morgan MT, et al. Bacterial communities in women with bacterial vaginosis: high resolution phylogenetic analyses reveal relationships of microbiota to clinical criteria. *PLoS One.* 2012;7(6):e37818.

58. Golob JL, Pergam SA, Srinivasan S, et al. Stool Microbiota at Neutrophil Recovery Is Predictive for Severe Acute Graft vs Host Disease After Hematopoietic Cell Transplantation. *Clin Infect Dis.* Nov 29 2017;65(12):1984-1991.

59. Matsen FA, Kodner RB, Armbrust EV. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics.* Oct 30 2010;11:538.

60. Lozupone CA, Knight R. Species divergence and the measurement of microbial diversity. *FEMS Microbiol Rev.* Jul 2008;32(4):557-578.

61. Navas-Molina JA, Peralta-Sanchez JM, Gonzalez A, et al. Advancing our understanding of the human microbiome using QIIME. *Methods Enzymol.* 2013;531:371-444.

62. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. *Genome Biol.* Jun 24 2011;12(6):R60.

63. Afgan E, Baker D, Batut B, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* Jul 2 2018;46(W1):W537-W544.

64. Doyle R, Gondwe A, Fan YM, et al. A Lactobacillus-Deficient Vaginal Microbiota Dominates Postpartum Women in Rural Malawi. *Appl Environ Microbiol.* Mar 15 2018;84(6).

65. MacIntyre DA, Chandiramani M, Lee YS, et al. The vaginal microbiome during pregnancy and the postpartum period in a European population. *Sci Rep.* Mar 11 2015;5:8988.

66. Mirmonsef P, Hotton AL, Gilbert D, et al. Glycogen Levels in Undiluted Genital Fluid and Their Relationship to Vaginal pH, Estrogen, and Progesterone. *PLoS One.* 2016;11(4):e0153553.

67. Chaban B, Links MG, Jayaprakash TP, et al. Characterization of the vaginal microbiota of healthy Canadian women through the menstrual cycle. *Microbiome.* 2014;2:23.

68. Achilles SL, Hillier SL. The complexity of contraceptives: understanding their impact on genital immune cells and vaginal microbiota. *AIDS.* Oct 2013;27 Suppl 1:S5-15.

69. *Trends in Contraceptive Use Worldwide 2015*: United Nations;2015.

70. Brind J, Condly SJ, Mosher SW, Morse AR, Kimball J. Risk of HIV Infection in Depot-Medroxyprogesterone Acetate (DMPA) Users: A Systematic Review and Meta-analysis. *Issues Law Med.* Autumn 2015;30(2):129-139.

71. Morrison CS, Chen PL, Kwok C, et al. Hormonal contraception and the risk of HIV acquisition: an individual participant data meta-analysis. *PLoS Med.* Jan 2015;12(1):e1001778.

72. Ralph LJ, McCoy SI, Shiu K, Padian NS. Hormonal contraceptive use and women's risk of HIV acquisition: a meta-analysis of observational studies. *Lancet Infect Dis.* Feb 2015;15(2):181-189.

73. Passmore J, Williams B. Role of vaginal microbiota in genital inflammation and enhancing HIV transmission. *International AIDS Conference*. Durban, South Africa2016.

74. Srinivasan S, Richardson BA, Wallis J, et al. Vaginal Microbiota and HIV Acquisition Risk Among African Women. *CROI*. Boston, MA, USA2018.

75. Nanda K, Callahan R, Taylor D, et al. Medroxyprogesterone acetate levels among Kenyan women using depot medroxyprogesterone acetate in the FEM-PrEP trial. *Contraception.* Jul 2016;94(1):40-47.

76. Fredricks DN, Fiedler TL, Thomas KK, Mitchell CM, Marrazzo JM. Changes in vaginal bacterial concentrations with intravaginal metronidazole therapy for bacterial vaginosis as assessed by quantitative PCR. *J Clin Microbiol.* Mar 2009;47(3):721-726.

77. Srinivasan S, Morgan MT, Fiedler TL, et al. Metabolic signatures of bacterial vaginosis. *MBio.* Apr 14 2015;6(2).

78. Fredricks DN, Fiedler TL, Thomas KK, Oakley BB, Marrazzo JM. Targeted PCR for detection of vaginal bacteria associated with bacterial vaginosis. *J Clin Microbiol.* Oct 2007;45(10):3270-3276.

79. Austin PC, Steyerberg EW. The number of subjects per variable required in linear regression analyses. *J Clin Epidemiol.* Jun 2015;68(6):627-636.

80. Thomson KA, Hughes J, Baeten JM, et al. Increased Risk of HIV Acquisition Among Women Throughout Pregnancy and During the Postpartum Period: A Prospective Per-Coital-Act Analysis Among Women With HIV-Infected Partners. *J Infect Dis.* Jun 5 2018;218(1):16-25.

81. Ahmed A, Earl J, Retchless A, et al. Comparative genomic analyses of 17 clinical isolates of Gardnerella vaginalis provide evidence of multiple genetically isolated clades consistent with subspeciation into genovars. *J Bacteriol.* Aug 2012;194(15):3922-3937.

82. Santiago GL, Deschaght P, El Aila N, et al. Gardnerella vaginalis comprises three distinct genotypes of which only two produce sialidase. *Am J Obstet Gynecol.* May 2011;204(5):450 e451-457.

83. Schellenberg JJ, Patterson MH, Hill JE. Gardnerella vaginalis diversity and ecology in relation to vaginal symptoms. *Res Microbiol.* Nov - Dec 2017;168(9-10):837-844.

84. Weinstock GM. Genomic approaches to studying the human microbiota. *Nature.* Sep 13 2012;489(7415):250-256.

85. McClelland RS, Richardson BA, Cherutich P, et al. A 15-year study of the impact of community antiretroviral therapy coverage on HIV incidence in Kenyan female sex workers. *AIDS.* Nov 2015;29(17):2279-2286.

86. Kinuthia J, Drake AL, Matemo D, et al. HIV acquisition during pregnancy and postpartum is associated with genital infections and partnership characteristics. *AIDS.* Sep 24 2015;29(15):2025-2033.

87. Celum C, Wald A, Lingappa JR, et al. Acyclovir and transmission of HIV-1 from persons infected with HIV-1 and HSV-2. *N Engl J Med.* Feb 4 2010;362(5):427-439.

88. Baeten JM, Donnell D, Ndase P, et al. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. *N Engl J Med.* Aug 2 2012;367(5):399-410.

89. Cheung YB, Ma X, Lam KF, Li J, Milligan P. Bias control in the analysis of case-control studies with incidence density sampling. *Int J Epidemiol.* Jun 17 2019.

90. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological).* 1996;58(1):267-288.

91. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology).* 2011;73(3):273-282.

92. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol.* Sep 2010;5(9):1315-1316.

93. Ahn J, Yang L, Paster BJ, et al. Oral microbiome profiles: 16S rRNA pyrosequencing and microarray assay comparison. *PLoS One.* 2011;6(7):e22788.

94. Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO. Development of the human infant intestinal microbiota. *PLoS Biol.* Jul 2007;5(7):e177.

95. Claesson MJ, O'Sullivan O, Wang Q, et al. Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS One.* Aug 20 2009;4(8):e6669.

96. Smith BC, McAndrew T, Chen Z, et al. The cervical microbiome over 7 years and a comparison of methodologies for its characterization. *PLoS One.* 2012;7(7):e40425.