Characterization of Vaginal Microbial Community Dynamics in the Pathogenesis of Incident Bacterial Vaginosis, a Pilot Study

Jacob H. Elnaggar, BS, John W. Lammons, BS,* Christopher M. Taylor, PhD,* Evelyn Toh, PhD,† Caleb M. Ardizzone, BS,* Amy Dong, HSD,† Kristal J. Aaron, DrPH, MSPH,‡ Meng Luo, PhD,* Ashutosh Tamhane, MD, PhD,§¶ Elliot J. Lefkowitz, PhD,¶‖ Alison J. Quayle, PhD,* David E. Nelson, PhD,† and Christina A. Muzny, MD, MSPH‡*

Background: Despite more than 60 years of research, the etiology of bacterial vaginosis (BV) remains controversial. In this pilot study, we used shotgun metagenomic sequencing to characterize vaginal microbial community changes before the development of incident BV (iBV).

Methods: A cohort of African American women with a baseline healthy vaginal microbiome (no Amsel criteria, Nugent score 0–3 with no *Gardnerella vaginalis* morphotypes) were followed for 90 days with daily self-collected vaginal specimens for iBV (≥2 consecutive days of a Nugent score of 7–10). Shotgun metagenomic sequencing was performed on select vaginal specimens from 4 women, every other day for 12 days before iBV diagnosis. Sequencing data were analyzed through Kraken2 and bioBakery 3 workflows, and specimens were classified into community state types. Quantitative polymerase chain reaction was performed to compare the correlation of read counts with bacterial abundance.

Results: Common BV-associated bacteria such as *G. vaginalis*, *Prevotella bivia*, and *Fannyhessea vaginae* were increasingly identified in the participants before iBV. Linear modeling indicated significant increases in *G. vaginalis* and *F. vaginae* relative abundance before iBV, whereas the relative abundance of *Lactobacillus* species declined over time. The *Lactobacillus* species decline correlated with the presence of *Lactobacillus* phages. We observed enrichment in bacterial adhesion factor genes on days before iBV. There were also significant correlations between bacterial read counts and abundances measured by quantitative polymerase chain reaction.

Conclusions: This pilot study characterizes vaginal community dynamics before iBV and identifies key bacterial taxa and mechanisms potentially involved in the pathogenesis of iBV.

Bacterial vaginosis (BV) is the most common cause of vaginal discharge in women,[1] and this dysbiosis is strongly associated with adverse birth outcomes,[2] pelvic inflammatory disease,[3] and an increased risk of acquisition and transmission of sexually transmitted infections (STIs), including HIV.[4–6] The prevalence of BV is higher among African American (AA) women (51.4%) compared with White women (23.2%),[7] and the annual economic burden of prevalent BV is estimated to be 4.8 billion US dollars.[8] A large body of epidemiologic data suggests that BV is an STI, with a hypothesized incubation period of 4 to 7 days,[9,10] but its exact etiology remains unknown.

Despite more than 60 years of research, the specific mechanisms of incident BV (iBV) pathogenesis remain unclear. The vaginal microbiome of women with BV is characterized by reduced numbers of lactic acid and hydrogen peroxide–producing lactobacilli and dramatically increased numbers of facultative and strict anaerobic BV-associated bacteria (BVAB), including *Gardnerella vaginalis*, *Prevotella* species, *Fannyhessea vaginae,* and *Megasphaera* species, compared with the vaginal microbiome of women without BV.[6] It has been proposed that BVAB could be

From the *Department of Microbiology, Immunology, and Parasitology, Louisiana State University Health Sciences Center, New Orleans, LA; †Department of Microbiology, Indiana University School of Medicine, Indianapolis, IN; and ‡Divisions of Infectious Diseases and §Nephrology, ¶Center for Clinical and Translational Sciences, and ‖Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL

Amy Dong is now with Harvard University, Cambridge, MA.

J.H.E. and J.W.L. contributed equally to the work.

Acknowledgments: The authors thank Shaunte Fisher, Laurie Hillsman, Sanquetta McClendon, and Joy Lewis for assisting with participant recruitment for this study. They also thank Jane Schwebke, MD; Hanne Harbison, MHS, MSN, CRNP; Saralyn Richter, RN; Rhonda Whidden, RN; Meghan Whitfield, CRNP; Cynthia Poore, CRNP; and Christen Press, CRNP, for assisting with participant enrollment, in addition to Charles Rivers, PhD, MSPH; Keonte Graves, MS; and Cheri Aycock for assisting with Nugent score reading. This study was accepted as a poster presentation at Anaerobe 2020, which was later canceled because of the COVID-19 pandemic. In addition, portions of this study were presented as an oral presentation at the 2021 UAB Microbiome Symposium on January 13, 2021, in Birmingham, AL; an oral presentation at the 2022 APSA Northeast Regional Conference on January 16, 2022, in Pittsburgh, PA; and a poster presentation at the 2022 Infectious Diseases Society for Obstetrics and Gynecology meeting on August 4, 2022, in Boston, MA.

Conflict of Interest and Sources of Funding: C.A.M. is a consultant for Scynexis, Cepheid, and BioNTech. She has received research funding support from National Institutes of Health/National Institute of Allergy and Infectious Diseases, Lupin Pharmaceuticals, Abbott Molecular, and Gilead, as well as honoraria from Visby, Elsevier, Abbott Molecular, Scynexis, Roche, and Cepheid. This work was supported by the National Institute of Allergy and Infectious Diseases (grants K23AI106957 and R01AI146065 to C.A.M. and R01AI118860 to A.J.Q. and C.M.T.), National Science Foundation (grant 2018936 to C.M.T.), National Center for Advancing Translational Sciences (grant UL1TR003096 to the University of Alabama–Birmingham Center for Clinical and Translational Science), and an American Sexually Transmitted Diseases Association Small Project Assistance award (to J.H.E.).

Correspondence: Christopher M. Taylor, PhD, Louisiana State University Health Sciences Center, 533 Bolivar St, New Orleans, LA. E-mail: ctay15@lsuhsc.edu.

Received for publication January 25, 2023, and accepted April 6, 2023.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text, and links to the digital files are provided in the HTML text of this article on the journal's Web site (http://www.stdjournal.com).

DOI: 10.1097/OLQ.0000000000001821

introduced via sexual transmission.[11–13] Alternatively, the proliferation of BVAB could be triggered by the loss of protective vaginal *Lactobacillus* species due to menses, sexual activity, *Lactobacillus* phage, and/or other inciting factors.[14–17] An interplay between these hypotheses could also be occurring, and sexual transmission of BVAB along with loss of protective lactobacilli could represent successive steps in developing BV.[18]

In our prior study, we characterized the vaginal microbiota of a cohort of AA women who have sex with women (WSW) over a 90-day period using daily self-collected vaginal specimens and 16S rRNA gene sequencing.[19] We determined that the relative abundance of *Lactobacillus crispatus* significantly decreased in participants who developed iBV (Nugent score 7–10 on ≥2 consecutive days) starting 14 days before diagnosis, compared with age- and race-matched women who did not. Conversely, the relative abundance of several key BVAB significantly increased 3 to 4 days before the development of iBV (*Prevotella bivia*, *G. vaginalis*) or on the day of iBV (*F. vaginae*, *Megasphaera* type I).[19] Because 16S rRNA gene sequencing only measures the relative abundance of bacteria present in a specimen, it was unclear from these data whether absolute increases or decreases of these BVAB occurred before iBV.

Although 16S rRNA gene sequencing has provided critical insights into the composition of vaginal bacterial communities during health and disease,[20] this approach has limitations, as mentioned previously. Shotgun metagenomic sequencing provides a more comprehensive insight into the composition of the vaginal microbiota, including genomic and functional data on all bacterial and nonbacterial DNA present in the vaginal microenvironment.[20] It can often yield a finer phylogenetic resolution of the bacterial taxa and, we hypothesize, a more accurate measurement of absolute abundance.

In this study, we analyzed longitudinal vaginal specimens from a subset of AA WSW who developed iBV from our prior study using shotgun metagenomic sequencing to further study changes in the vaginal microbiota before iBV to better understand its pathogenesis.

MATERIALS AND METHODS

Study Population and Sampling

This substudy and the parent study were approved by the University of Alabama at Birmingham Institutional Review Board (IRB; Protocol F131127001),[19] the Jefferson County Department of Health Research Review Committee (Protocol 2015-03), and the Louisiana State University IRB (Protocol 8738). This study was exempt by the Indiana University IRB. In the parent study, 42 AA nonpregnant WSW with baseline normal vaginal microbiota (no Amsel criteria[21] and a normal Nugent score[22] of 0–3 with no *G. vaginalis* morphotypes), no oral or intravaginal antibiotic use within the past 14 days, and no current STIs (gonorrhea, chlamydia, and trichomoniasis) were enrolled at the Jefferson County Department of Health Sexual Health Clinic in Birmingham, Alabama.[19] These women provided daily self-collected vaginal specimens for 90 days or until the development of iBV (Nugent score 7–10 on ≥2 consecutive days). Participants completed a study enrollment questionnaire to obtain sociodemographic information as well as medical and sexual history data. Participants also filled out daily diaries about their sexual partners and menses throughout the study period.[19]

Based on previous 16S rRNA gene sequencing (Fig. S1, Heatmap of 16S rRNA gene sequencing from the 4 participants, http://links.lww.com/OLQ/A940),[19] we selected a subset of stored vaginal specimens from 4 participants who developed iBV for the purposes of this pilot study. We selected specimens from the first day of iBV diagnosis (day 0), along with every other day for the 10 days before iBV (days −2 to −10). If a vaginal specimen was unavailable (i.e., day −8 for K12), the next available specimen closest in time was chosen.

Shotgun Metagenomic Sequencing

DNA was isolated, as previously described[19]; sent to the Department of Microbiology and Immunology at Indiana University; and prepared for shotgun metagenomic sequencing. The NexteraXT library prep kit (Illumina, San Diego, CA) was used, and samples were paired-end sequenced (2 × 150 bp) on a HiSeq4000. Raw fastq files were used for bioinformatics analysis. Sequences were deposited in the NCBI Sequence Read Archive (PRJNA707585) after removing human reads (Table S1A, Removal of human reads, http://links.lww.com/OLQ/A941).

Taxonomic Assignment and Processing

We used multiple processing and classification tools to rigorously classify and analyze the sequence reads from this pilot study. First, Kraken2 v2.0.9-beta was used to map the sequences against the MaxiKraken database (Loman Lab, https://lomanlab.github.io/mockcommunity/mc_databases.html).[23] Organisms assigned less than 10 reads by Kraken2, and potential reagent contaminants were removed from further analysis (Table S1B, Contamination removal, http://links.lww.com/OLQ/A941). After analysis with Kraken2, further analysis was carried out using the bioBakery 3 suite of tools.[24] VIRGO was used as a third method for taxonomic classification.[25] Relative abundance was obtained by dividing each microorganism's read count by the total read count in each specimen to adjust for sequencing variation. Maximal reads were chosen because of the hypothesized significant increase in bacterial reads on the day of iBV that may skew filtering if averaging was performed.[6,11]

Metagenomic Bacterial Analysis

Bacterial strain composition was determined using Panphlan 3.0 to map shotgun reads to species pangenomes and reference strain gene sets.[26] Bray-Curtis dissimilarity distances were calculated from gene sets generated from Panphlan 3. Results were visualized using Principal Coordinates Analysis (PCoA) and clustered using k-means. KEGG overrepresentation analysis and virulence factor (VF) identification was performed using cluster profiler 4.0 and ShortBRED, respectively.[27,28] Community state types (CSTs) were determined with VALENCIA.[29]

Metagenomic Phage Analysis

To investigate *Lactobacillus* phage origin and activity as part of this analysis, we performed metagenomic assembly using the ATLAS workflow, and phage contigs were identified using VIBRANT.[30,31s] Further details of these methods are described in Supplemental Methods, http://links.lww.com/OLQ/A942.

Bacterial Abundance Measurement

We used quantitative polymerase chain reaction (qPCR) assays to compare the correlation between sequence reads and bacterial abundance. The proportion of bacterial reads to total reads was compared with the total bacterial load (16S rRNA copy number) using universal V6 primers, as described.[32s] Based on their predominant relative abundance in our samples, we compared the log-scaled relative abundance of *G. vaginalis* and *Lactobacillus* species from Kraken2, Metaphlan 3, and VIRGO to their 16S copy numbers using specific V6 primers. Additional qPCR information is listed in Table S1C (qPCR information, http://links.lww.com/OLQ/A941).

Statistical Analysis

All visualizations were generated in R v4.2.1 using the package ggplot2.[33s,34s] Statistical analysis used the R package MaAslin2,[35s] which specializes in multivariable associations in microbiome data using linear mixed models. We implemented this method to account for correlated Nugent scores, temporal metrics, phages, and 16S rRNA gene copy numbers to species abundance. To adjust for multiple testing and therefore inflated type I error rate, we adjusted *P* values using false discovery rate, and a *P* value of less than 0.10 (2-tailed) was considered statistically significant.[36s]

RESULTS

Characteristics of the Study Population

Basic demographic, sexual history, and tobacco use data for the 4 participants with iBV included in this substudy (K12, K18, K19, K20) are listed in Table 1. A total of 24 vaginal specimens from these 4 participants were sequenced (6 specimens/woman). Figure 1A indicates relevant metadata for each woman, along with an indication of which days were sequenced via 16S rRNA gene sequencing in the parent study and shotgun metagenomic sequencing in this pilot study.[19]

Microbial Composition of the Vaginal Specimens

More than 85% of the shotgun metagenomic sequences were human; bacteria accounted for most nonhuman sequences (Fig. 1B). There was a significant increase in reads originating from bacterial microorganisms on the day of iBV in 3 of the 4 women, with most of the remaining sequences corresponding to fungi (Fig. 1C). The distribution of viral sequences is displayed in Figure 1D. An increase in viral reads occurred in K20 corresponding to human gammaherpesvirus 4 (Epstein-Barr virus) before iBV, which may be a marker of cellular infiltrate (B cells).[37s] Overall, we found more precise assignment of species using shotgun metagenomic sequencing compared with our previous 16S analysis (Fig. S2, 16S rRNA gene sequencing and shotgun metagenomics taxonomic comparison, http://links.lww.com/OLQ/A943).[19]

Ordination of Bray-Curtis dissimilarity distances from sample species relative abundance reflects the microbial community dynamics occurring during the onset of iBV for each participant (Fig. 1E). The difference in Bray-Curtis distance between consecutively collected samples is plotted in (Fig. S3A, Comparisons of distances between longitudinal specimens, http://links.lww.com/OLQ/A944). Categorization and scores of CST are reported in Table S1D (CST scores, http://links.lww.com/OLQ/A941). In 3 of the 4 participants sampled, the greatest change in distance occurred in communities labeled as CST III, *Lactobacillus iners,* dominated communities. This observation aligns with previous findings suggesting that *L. iners* communities are transitional.[38s]

The most abundant microorganisms are displayed in heatmaps (Figs. 2A, B), ordered by maximum read count across all samples. Analysis with Kraken2 (Fig. 2A) revealed that the highest overall relative abundance was attributed to *G. vaginalis*, *F. vaginae*, and *P. bivia* on the day of iBV. Taxonomic classification using the vaginal microbiome-specific taxonomic classification tool VIRGO (Fig. S4, Taxonomic analysis using VIRGO, http://links.lww.com/OLQ/A945) generated similar results to Metaphlan 3.0 (Fig. 2B); however, a greater relative abundance was attributed to *Lactobacillus* species. Overall, bacterial taxa detected before and on the day of iBV matched previously reported findings.[19]

The relative abundance of several common BVAB increased leading up to iBV, whereas *Lactobacillus* species declined. To confirm this observation, we generated linear models correlating Nugent scores and temporal metrics to species relative abundance using MaAslin2 (Fig. S5A, B, MaAslin2 results, http://links.lww.com/OLQ/A946; Table S1E, MaAslin2 results, http://links.lww.com/OLQ/A941). We found statistically significant associations with *F. vaginae* ($P = 0.04$) and *G. vaginalis* ($P = 0.07$) relative abundance increasing as the day of iBV approached. *P. bivia* relative abundance significantly increased before iBV ($P = 0.08$) in subject K18 using linear models. We also found significant decreases in the relative abundance of *L. iners* ($P = 0.09$), *L. crispatus* ($P = 0.01$), and *Lactobacillus jensenii* ($P = 0.07$) across all specimens as well as *Lactobacillus gasseri* ($P = 0.08$) in K18 on days leading up to iBV. When categorizing Nugent scores as low (0–3), intermediate (4–6), and high (7–10), we found that the relative abundance of *F. vaginae* was only >1% on days of high Nugent scores, suggesting that it functions as a mediator of iBV later in onset.[11]

Functional Annotations Before iBV

Functional annotations were performed using HUMAnN 3.0 (Fig. 3A). Samples were clustered by CST. The pathway with the highest relative abundance on the day of iBV was the glycogen degradation II pathway. The relative abundance of this pathway was found to be significantly higher on the day of iBV compared with days leading up to iBV ($P = 0.01$). *G. vaginalis* was found to be the principal contributor to this pathway (Table S1F, Species contributions to pathways, http://links.lww.com/OLQ/A941).[39s] In addition, the pathway 5-aminoimidazole ribonucleotide biosynthesis was significantly enriched in samples with iBV CST compared with type I and III CSTs ($P = 0.01$).

To further assess functional changes, reads were mapped to VF libraries. *L. iners*– and *G. vaginalis*–specific VFs, inerolysin, and vaginolysin, respectively, showed unique gene abundance patterns, which closely mirrored the relative abundance of each species. In addition, genes associated with chaperones and adhesion proteins were more abundant in samples leading up to the day of iBV.[40s,41s] This was particularly clear in samples from participant

TABLE 1. Characteristics of the Study Participants

Characteristic	K12	K18	K19	K20
Age, y	19	34	39	28
Race	AA	AA	AA	AA
Education	Some college	Some college	Some college	Less than high school
Tobacco use, past 30 d	No	Yes	Yes	Yes
Douching, past 30 d	No	No	No	No
Vaginal lubricant use, past 3 mo	No	No	No	No
Self-reported STI history*	Yes	No	Yes	Yes
Contraception use	Past	Never	Never	Past

*Includes history of *Trichomonas vaginalis*, *Neisseria gonorrhoeae*, *Chlamydia trachomatis*, genital herpes, and/or syphilis infection.

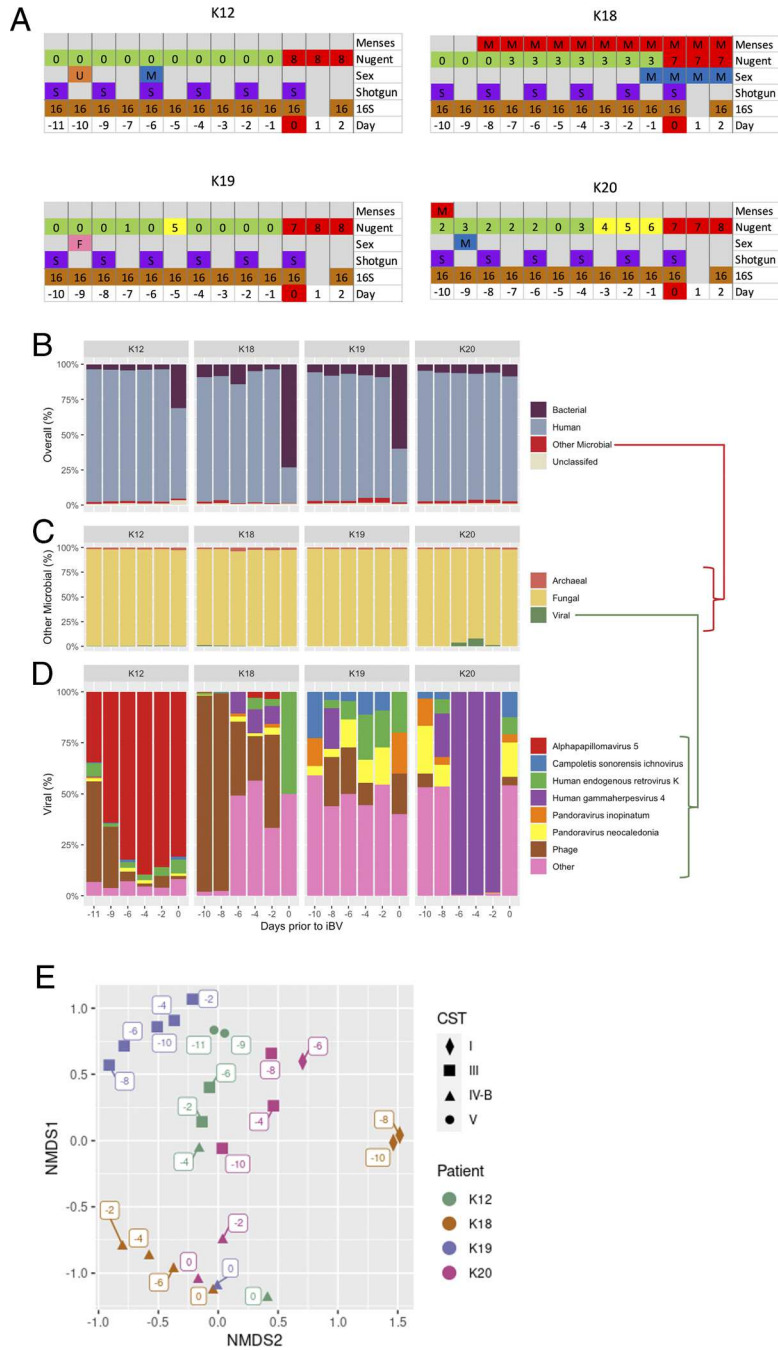

Figure 1. Composition of cohort and specimens. A, Schematic of the sampling days for each woman with associated metadata. Days before iBV (red) are indicated for each woman in the bottom row. The subset of specimens that were sequenced via shotgun metagenomics is indicated with S (purple) in the shotgun row. Days of reported sexual activity are indicated with a female (F; pink), male (M; blue), or unknown (U; orange) in the Sex row. Nugent scores (0–10), color coded as normal = 0–3 (green), intermediate = 4–6 (yellow), or BV = 7–10 (red), are indicated in the Nugent row. Days that menses were reported are indicated with an M (red) in the top row. B–D, Composition of taxonomic reads found across all 24 specimens. Stacked bar charts displaying the percentage of each taxonomic classification among specific breakdowns. Specific groups are then further broken down as depicted with brackets on the right. B contains the largest categories of organisms, including human, C breaks down the other microbial reads, D breaks down the viral reads. Classification was performed using Kraken2 with the maxiKraken database. E, NMDS plot of sample Bray-Curtis dissimilarity distances generated from the taxonomic composition. Figure 1 can be viewed online in color at www.stdjournal.com.

K19. Row normalized VF abundance is shown in Fig. S3B (Heatmap of VF abundance, http://links.lww.com/OLQ/A944).

G. vaginalis is thought to exist in both commensal and pathogenic clades, as it is detected in women with and without BV.[42s] To better characterize *G. vaginalis* in the development of iBV, we investigated the genetic composition of strains found in our specimens. *G. vaginalis* strains were identified in 19 of 24 samples. Strain similarity was visualized using a PCoA plot

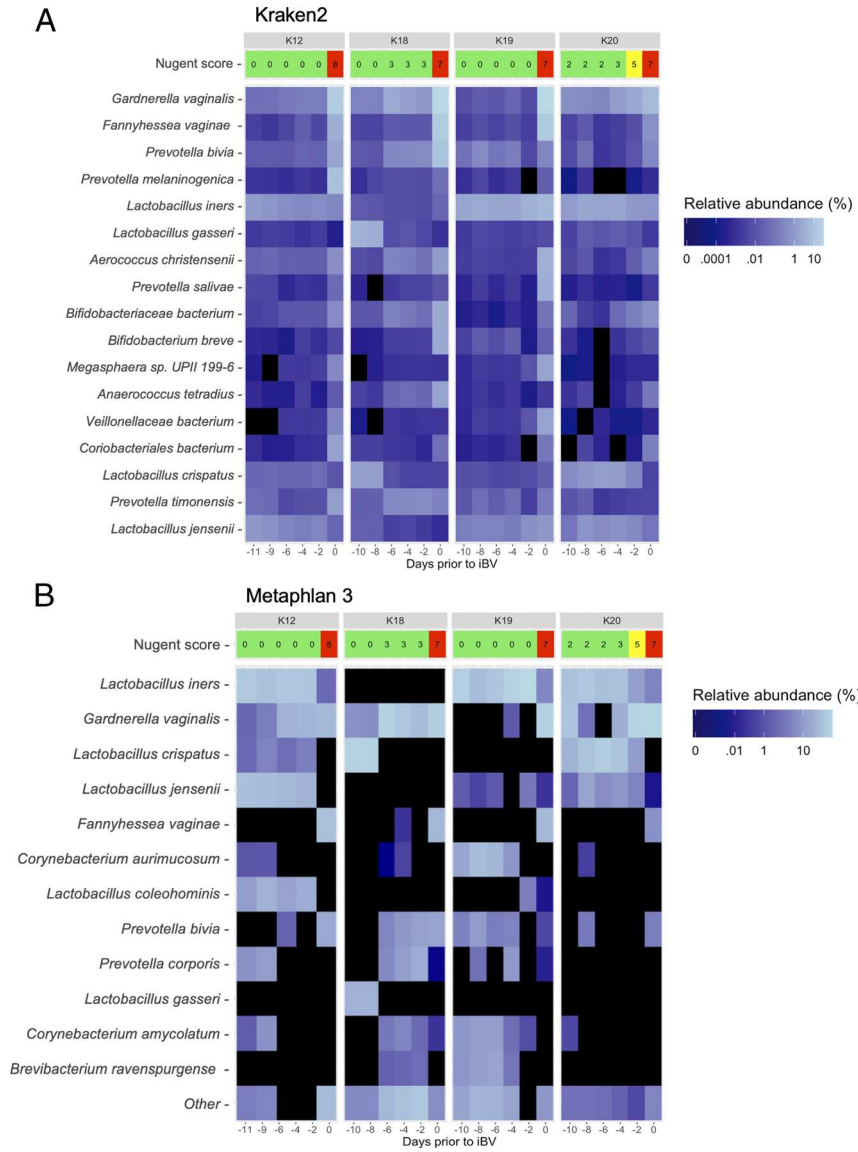

Figure 2. Bacterial taxonomic classification. A and B, Bacterial heatmaps for K12, K18, K19, and K20. Heatmap for the top 25% of bacterial microorganisms found in the 4 participants. Identified bacteria are ordered by their log-normalized maximum read count across all specimens. Each participant is represented in a vertical block. The Nugent score is displayed at the top and color coded as normal = 0–3 (green), intermediate = 4–6 (yellow), or BV = 7–10 (red). Classification was performed using (A) Kraken2 or (B) Metaphlan 3. Figure 2 can be viewed online in color at www.stdjournal.com.

(Fig. 3B). The dominant strain of *G. vaginalis* in 15 of the 19 samples clustered with reference genomes associated with strains previously identified to be from clade 1.[43s] Despite identifying strains from multiple clades, we did not find a difference in strain pathway enrichment on days leading up to iBV compared with the day of iBV.

Lactobacillus Phages in iBV Pathogenesis

We sought to better characterize the bacteriophage population because of the presence of phages in some quantity in all participants before iBV (Fig. 1D). Although phage reads made up a small proportion of total microbial reads, they made up a larger proportion of viral reads (~10%). Analysis with VIBRANT indicated that the majority of the assembled *Lactobacillus* phage contigs did not also contain bacterial DNA (Fig. S6, http://links.lww.com/OLQ/A947; Table S1G, *Lactobacillus* phage results, http://links.lww.com/OLQ/A941).[31s] The total *Lactobacillus* phage relative abundance decreased leading up to iBV ($P = 0.01$; Fig. S6, http://links.lww.com/OLQ/A947; Fig. S5C, MaAslin2 results, http://links.lww.com/OLQ/A946; Table S1E, MaAslin2 results, http://links.lww.com/OLQ/A941). In addition, total *Lactobacillus* phage relative abundance positively correlated with *L. crispatus* ($P = 0.09$), and both decreased before iBV (Fig. 4). Phage contig coverage was found to be significantly higher than associated host *Lactobacillus* coverage in 6 samples on days leading up to iBV ($P < 0.01$), indicating that *Lactobacillus* phages could be active in these samples (Table S1J, Phage coverage and activity, http://links.lww.com/OLQ/A941).

Estimating Abundance Using qPCR

The ratio of bacterial reads to total sequencing reads was hypothesized to be an estimate of absolute bacterial abundance

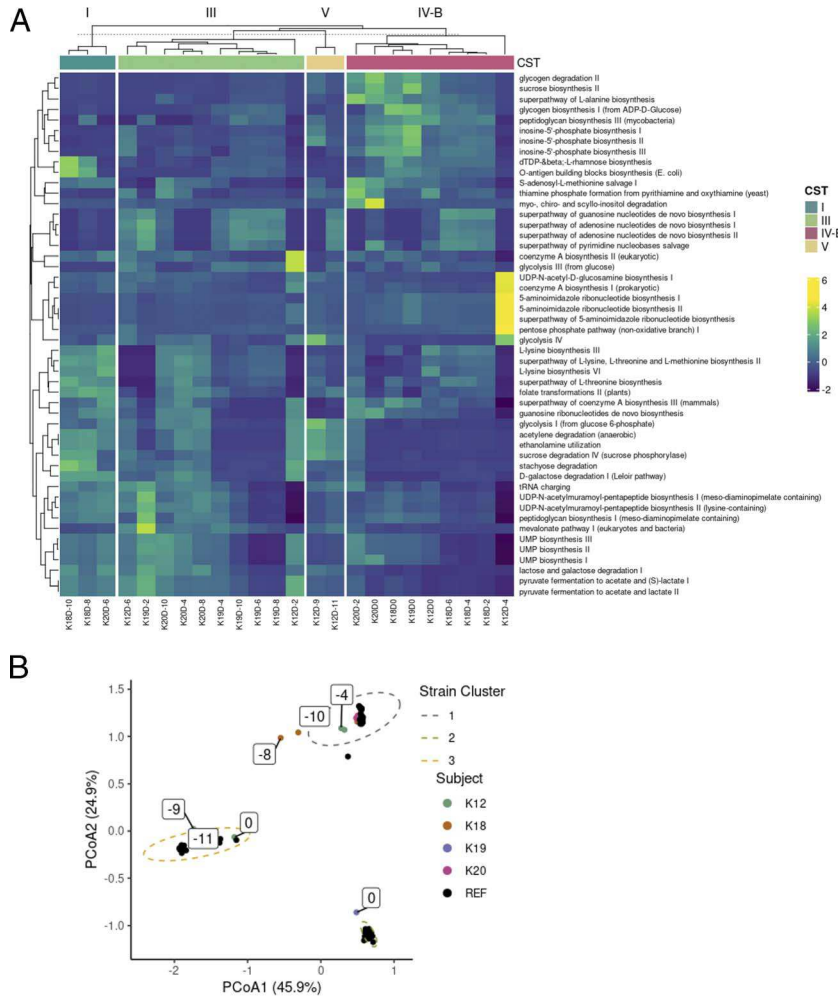

Figure 3. Functional information before iBV. A, Heatmap of Metacyc and KEGG pathways abundance generated from Metaphlan 3. Heatmap of the relative abundance of the top 50 most abundant pathways generated from HUMAnN 3.0. Values are row normalized, and columns are clustered by sample CST. B, PCoA of *G. vaginalis* strain similarity. Bray-Curtis dissimilarity distances are generated from strain gene sets and plotted using PCoA. Samples are clustered using k-means. Reference strains are plotted with strains identified in participant samples. A 95% confidence interval was plotted for each cluster. Cluster 1 reflects *G. vaginalis* strains found in clade 1, cluster 2 contains G. *vaginalis* from clade 2, and cluster 3 contains *G. vaginalis* strains from clades 3 and 4. Figure 3 can be viewed online in color at www.stdjournal.com.

in the specimens.[44s] The estimated abundance of bacterial reads increased on the day of iBV for all 4 participants (Fig. 2B). In 3 of the 4 participants (K12, K18, K19), this increase was over 10-fold compared with the prior 5 vaginal specimens. The fourth participant, K20, showed a nearly 2-fold increase on the day of iBV. To verify that these reads could estimate absolute abundance, qPCR assays measuring total bacterial load (16S rRNA copy number), as well as *G. vaginalis* and *Lactobacillus* species, were performed. V6-specific 16S rRNA copy numbers were compared with the corresponding read counts from Kraken2, Metaphlan 3, and VIRGO using MaAslin2 (Fig. S5D, MaAslin2 results, http://links.lww.com/OLQ/A946; Table S1E, http://links.lww.com/OLQ/A941). We observed a strong correlation between the total bacterial load to the proportion of bacterial reads to total sequenced reads (including human reads, $P = 0.001$). Comparing the qPCR-derived copies with read counts for *Lactobacillus* species and *G. vaginalis*, we observed the strongest correlation using Kraken2 ($P = 4E-9$, $1.5E-5$) followed by Metaphlan 3 ($P = 1E-6$, $7E-4$) and VIRGO ($P = 0.05$, 0.04). Overall, these results indicate that estimations of bacterial absolute abundance can be made based on the ratio of bacterial to total sequencing reads.

DISCUSSION

Discovering the etiology of iBV remains an active area of research, and the causal pathway leading to iBV is yet to be identified.[11,12,45s] It is unknown whether vaginal lactobacilli decrease first in iBV pathogenesis, allowing BVAB to proliferate, or whether BVAB are introduced first, suppressing vaginal lactobacilli.[11] Various factors leading to the depletion of vaginal *Lactobacillus* species in iBV pathogenesis have been studied, including but not limited to douching,[46s] spermicides (e.g., Nonoxynol-9),[47s] and *Lactobacillus* phages.[14–17]

Our findings in this pilot study indicate that *G. vaginalis* increases in relative abundance on days leading up to iBV, whereas *F. vaginae* increases in relative abundance on the day of iBV. There are data that *P. bivia* is involved in the early stages of BV biofilm formation and iBV pathogenesis.[11,19,48s,49s] Although *P. bivia* was observed in all participants in some quantity in this study, we

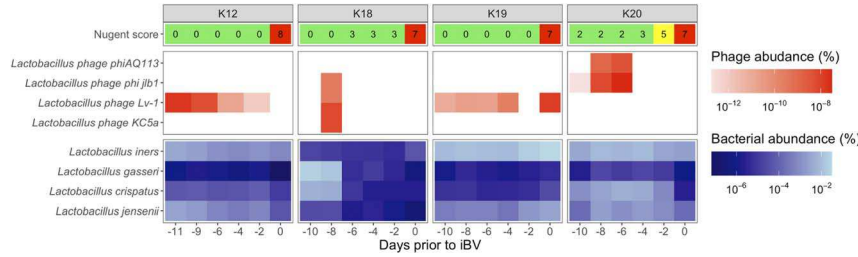

Figure 4. Lactobacillus phages before iBV. A, Lytic and lysogenic *Lactobacillus* phages. This figure displays all *Lactobacillus* phages identified using VIBRANT and Kraken2 (viral RefSeq database) and the corresponding log10 normalized relative abundance per specimen. The top displays the phages that were found to be lytic in red, and the bottom shows the lysogenic phages in blue. B, Heatmap of *Lactobacillus* phages. Heatmap for *Lactobacillus* phages identified to be lytic. The phages were chosen by having more than 500 reads in at least one specimen. Below displays the different *Lactobacillus* species for comparison. Bacterial reads were classified using Kraken2 with the maxiKraken database. Each participant is represented in a vertical block. The Nugent score is displayed at the top and color coded as normal = 0–3 (green), intermediate = 4–6 (yellow), or BV = 7–10 (red). Figure 4 can be viewed online in color at www.stdjournal.com.

observed a significant association between *P. bivia* abundance and iBV onset in only 1 of the 4 participants. This finding may be due to the small sample size of this study, limiting the ability to identify exact changes in bacterial abundance; future studies should include larger sample sizes. In addition, results of this study support the notion that *L. iners*–dominated communities are less resistant to rapid transitions.[38s]

Analysis of functional annotations showed enrichment in adhesion-related genes on days leading up to iBV. It is hypothesized that adhesion of *G. vaginalis* to the vaginal epithelium is the first step in BV biofilm formation.[11] This finding may reflect an increased abundance of genes involved in cell adhesion on days before iBV. Identifying adhesion proteins that accumulate before iBV could be important in determining proteins contributing to the pathogenesis of *G. vaginalis.* In addition, we observed an enrichment in glycogen degradation genes copies associated with *G. vaginalis,* including on the day of iBV, reflecting the macromolecules used at the time of iBV.

We identified *Lactobacillus* phages as far as 10 days before iBV in all 4 of the participants in this study. These phages correlated with a decrease in protective *Lactobacillus* species.[50s] A depletion of protective *Lactobacillus* species by phages may play a role in the pathogenesis of iBV, although this requires further study.[15] *Lactobacillus* phage phi jlb1 and KC5a are reported to target *L. gasseri* in vitro,[51s,52s] whereas *Lactobacillus* phage Lv-1 has been found to target *L. jensenii* in the vaginal exudate of a healthy, reproductive-aged woman.[53s] Phages targeting *L. crispatus* have been challenging to delineate as historical studies were performed on *Lactobacillus acidophilus* (targeted by KC5A), which was later reclassified into multiple species, including *L. crispatus*.[54s] A prior study of vaginal specimens from 39 reproductive-aged women identified 7 *Lactobacillus* phages from 37 *Lactobacillus* strains isolated from 27 women (19 with normal vaginal flora and 8 with BV based on Amsel criteria).[14] The rate of phage isolation was much lower in healthy women than in women with BV. Interestingly, phage ϕkc039, although lysogenic in one woman in which it was isolated from, was able to lyse lactobacilli in other women in this study. Although we did find enrichment in phage DNA compared with host DNA on days leading up to iBV, there is evidence that these phages could be lytic, although in vivo experiments will be needed to validate these findings. Additional studies on the potential role of sexual transmission of infectious *Lactobacillus* phages in iBV pathogenesis are needed.[15]

Strengths of this pilot study include increased resolution of taxonomic classifications for bacterial microorganisms compared with 16S rRNA gene sequencing and characterization of vaginal microorganisms longitudinally in a unique group of women with iBV. We used 3 widely used tools for shotgun metagenomic sequencing, Kraken2,[23] bioBakery 3,[24] and specifically for the vaginal microbiome, VIRGO.[25] In addition, we have shown that shotgun metagenomic sequencing estimates absolute bacterial abundance based on the proportion of bacterial to total sequencing reads; these results remained accurate for specific vaginal bacteria, *G. vaginalis* and *Lactobacillus* species.[44s]

Despite these strengths, this study has limitations. The cost of shotgun metagenomic sequencing is substantially higher than 16S rRNA gene sequencing. Thus, we only sequenced 4 participants resulting in our study being underpowered. Although we carried out statistical testing and found statistically significant differences, these findings can only estimate clinical significance. In addition, we were only able to sequence a smaller number of specimens (n = 24) over a limited period (every other day for the 10 days before iBV) than in our prior 16S study.[19] There could have been important changes occurring in the intervening days that we may have missed in this analysis. We also did not sequence specimens from participants who did not develop iBV to compare them with iBV women; this is an area of future research as we have stored specimens from control participants. In addition, the participants included in this study were AA WSW presenting to a sexual health clinic in the Southern United States, making it difficult to generalize these results to all women. However, BV is highly prevalent among this population, and this study could lead to targeted efforts to optimize sexual health in this community.[52s,53s] Also, phages (and other viruses) comprised only a small proportion of nonbacterial reads in this study (Fig. 1C), perhaps due to the self-collected nature of the sampling technique, which might not pick up all organisms equally.[55s] Future studies could consider clinician-collected specimens; however, this may not be feasible on a daily basis. To address these limitations, future well-powered studies with larger sample sizes, appropriate controls, and more frequent sequencing are needed.

In conclusion, this pilot study proposes interactions between specific vaginal organisms in the development of iBV. The results should be further explored in larger cohorts of women with and without iBV.

REFERENCES

1. Allsworth JE, Peipert JF. Prevalence of bacterial vaginosis: 2001–2004 National Health and Nutrition Examination Survey data. Obstet Gynecol 2007; 109:114–120.

2. Hillier SL, Nugent RP, Eschenbach DA, et al. Association between bacterial vaginosis and preterm delivery of a low-birth-weight infant. The Vaginal Infections and Prematurity Study Group. N Engl J Med 1995; 333:1737–1742.
3. Ness RB, Kip KE, Hillier SL, et al. A cluster analysis of bacterial vaginosis–associated microflora and pelvic inflammatory disease. Am J Epidemiol 2005; 162:585–590.
4. Cohen CR, Lingappa JR, Baeten JM, et al. Bacterial vaginosis associated with increased risk of female-to-male HIV-1 transmission: A prospective cohort analysis among African couples. PLoS Med 2012; 9:e1001251.
5. Eschenbach DA. Bacterial vaginosis and anaerobes in obstetric-gynecologic infection. Clin Infect Dis 1993; 16(Suppl 4):S282–S287.
6. Hillier S, Marrazzo J, Holmes K. Bacterial vaginosis. In: Holmes KK, Sparling PF, Mardh P-A, et al, eds. Sexually Transmitted Diseases. 4th ed. New York: McGraw-Hill, Health Profession Division; 2008:737–768.
7. Koumans EH, Sternberg M, Bruce C, et al. The prevalence of bacterial vaginosis in the United States, 2001–2004; associations with symptoms, sexual behaviors, and reproductive health. Sex Transm Dis 2007; 34: 864–869.
8. Peebles K, Velloza J, Balkus JE, et al. High global burden and costs of bacterial vaginosis: A systematic review and meta-analysis. Sex Transm Dis 2019; 46:304–311.
9. Muzny CA, Lensing SY, Aaron KJ, et al. Incubation period and risk factors support sexual transmission of bacterial vaginosis in women who have sex with women. Sex Transm Infect 2019; 95:511–515.
10. Bradshaw CS, Walker SM, Vodstrcil LA, et al. The influence of behaviors and relationships on the vaginal microbiota of women and their female partners: The WOW Health Study. J Infect Dis 2014; 209: 1562–1572.
11. Muzny CA, Taylor CM, Swords WE, et al. An updated conceptual model on the pathogenesis of bacterial vaginosis. J Infect Dis 2019; 220:1399–1405.
12. Srinivasan S, Fredricks DN. The human vaginal bacterial biota and bacterial vaginosis. Interdiscip Perspect Infect Dis 2008; 2008:750479.
13. Nelson DE, Van Der Pol B, Dong Q, et al. Characteristic male urine microbiomes associate with asymptomatic sexually transmitted infection. PloS One 2010; 5:e14116.
14. Pavlova SI, Kilic AO, Mou SM, et al. Phage infection in vaginal lactobacilli: An in vitro study. Infect Dis Obstet Gynecol 1997; 5: 36–44.
15. Blackwell AL. Vaginal bacterial phaginosis? Sex Transm Infect 1999; 75:352–353.
16. Tao L, Pavlova SI, Mou SM, et al. Analysis of lactobacillus products for phages and bacteriocins that inhibit vaginal lactobacilli. Infect Dis Obstet Gynecol 1997; 5:244–251.
17. Pavlova SI, Tao L. Induction of vaginal *Lactobacillus* phages by the cigarette smoke chemical benzo[*a*]pyrene diol epoxide. Mutat Res 2000; 466:57–62.
18. Turovskiy Y, Sutyak Noll K, Chikindas ML. The aetiology of bacterial vaginosis. J Appl Microbiol 2011; 110:1105–1128.
19. Muzny CA, Blanchard E, Taylor CM, et al. Identification of key bacteria involved in the induction of incident bacterial vaginosis: A prospective study. J Infect Dis 2018; 218:966–978.
20. Poretsky R, Rodriguez-R LM, Luo C, et al. Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. PloS One 2014; 9:e93827.
21. Amsel R, Totten PA, Spiegel CA, et al. Nonspecific vaginitis. Diagnostic criteria and microbial and epidemiologic associations. Am J Med 1983; 74:14–22.
22. Nugent RP, Krohn MA, Hillier SL. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. J Clin Microbiol 1991; 29:297–301.
23. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol 2019; 20:257.
24. Beghini F, McIver LJ, Blanco-Míguez A, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. ELife 2021; 10:e65088.
25. Ma B, France MT, Crabtree J, et al. A comprehensive non-redundant gene catalog reveals extensive within-community intraspecies diversity in the human vagina. Nat Commun 2020; 11:940.
26. Scholz M, Ward DV, Pasolli E, et al. Strain-level microbial epidemiology and population genomics from shotgun metagenomics. Nat Methods 2016; 13:435–438.
27. Wu T, Hu E, Xu S, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb) 2021; 2:100141.
28. Kaminski J, Gibson MK, Franzosa EA, et al. High-specificity targeted functional profiling in microbial communities with ShortBRED. PLoS Comput Biol 2015; 11:e1004557.
29. France MT, Ma B, Gajer P, et al. VALENCIA: A nearest centroid classification method for vaginal microbial communities based on composition. Microbiome 2020; 8:166.
30. Kieser S, Brown J, Zdobnov EM, et al. ATLAS: A Snakemake workflow for assembly, annotation, and genomic binning of metagenome sequence data. BMC Bioinformatics 2020; 21:257.

For further references, please see "Supplemental References," http://links.lww.com/OLQ/A948.