




Cervicovaginal Microbiome Composition Is Associated with Metabolic Profiles in Healthy Pregnancy

Andrew Oliver,^a Brandon LaMere,^b Claudia Weihe,^c Stephen Wandro,^d Karen L. Lindsay,^e Pathik D. Wadhwa,^{e,f} David A. Mills,^{g,h} David T. Pride,^{i,j} Oliver Fiehn,^k Trent Northen,^l Markus de Raad,^l Huiying Li,^m Jennifer B. H. Martiny,^c Susan Lynch,^b  Katrine Whiteson^a

^aDepartment of Molecular Biology and Biochemistry, University of California, Irvine

^bDivision of Gastroenterology, Department of Medicine, University of California, San Francisco, San Francisco

^cDepartment of Ecology and Evolutionary Biology, University of California, Irvine, Irvine, California, USA

^dCenter for Microbiome Innovation, University of California, San Diego, San Diego, California, USA

^eDepartment of Pediatrics and the Development, Health and Disease Research Program, College of Health Sciences, University of California, Irvine, Irvine, California, USA

^fDepartment of Psychiatry and Human Behavior, College of Health Sciences, University of California, Irvine, Irvine, California, USA

^gDepartment of Food Science and Technology, University of California, Davis, Davis, California, USA

^hDepartment of Viticulture & Enology, University of California, Davis, Davis, California, USA

ⁱDepartment of Pathology, University of California, San Diego, San Diego, California, USA

^jDepartment of Medicine, University of California, San Diego, San Diego, California, USA

^kWest Coast Metabolomics Center, University of California, Davis, Davis, California, USA

^lEnvironmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA

^mDepartment of Molecular and Medical Pharmacology, Crump Institute for Molecular Imaging, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California, USA

ABSTRACT Microbes and their metabolic products influence early-life immune and microbiome development, yet remain understudied during pregnancy. Vaginal microbial communities are typically dominated by one or a few well-adapted microbes which are able to survive in a narrow pH range and are adapted to live on host-derived carbon sources, likely sourced from glycogen and mucin present in the vaginal environment. We characterized the cervicovaginal microbiomes of 16 healthy women throughout the three trimesters of pregnancy. Additionally, we analyzed saliva and urine metabolomes using gas chromatography-time of flight mass spectrometry (GC-TOF MS) and liquid chromatography-tandem mass spectrometry (LC-MS/MS) lipidomics approaches for samples from mothers and their infants through the first year of life. Amplicon sequencing revealed most women had either a simple community with one highly abundant species of *Lactobacillus* or a more diverse community characterized by a high abundance of *Gardnerella*, as has also been previously described in several independent cohorts. Integrating GC-TOF MS and lipidomics data with amplicon sequencing, we found metabolites that distinctly associate with particular communities. For example, cervicovaginal microbial communities dominated by *Lactobacillus crispatus* have high mannitol levels, which is unexpected given the characterization of *L. crispatus* as a homofermentative *Lactobacillus* species. It may be that fluctuations in which *Lactobacillus* dominate a particular vaginal microbiome are dictated by the availability of host sugars, such as fructose, which is the most likely substrate being converted to mannitol. Overall, using a multi-“omic” approach, we begin to address the genetic and molecular means by which a particular vaginal microbiome becomes vulnerable to large changes in composition.

IMPORTANCE Humans have a unique vaginal microbiome compared to other mammals, characterized by low diversity and often dominated by *Lactobacillus* spp. Dramatic shifts in vaginal microbial communities sometimes contribute to the presence of a polymicrobial overgrowth condition called bacterial vaginosis (BV). However,

Citation Oliver A, LaMere B, Weihe C, Wandro S, Lindsay KL, Wadhwa PD, Mills DA, Pride DT, Fiehn O, Northen T, de Raad M, Li H, Martiny JBH, Lynch S, Whiteson K. 2020. Cervicovaginal microbiome composition is associated with metabolic profiles in healthy pregnancy. mBio 11:e01851-20. <https://doi.org/10.1128/mBio.01851-20>.

Editor Martin J. Blaser, Rutgers University

Copyright © 2020 Oliver et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Katrine Whiteson, katrine@uci.edu.

Received 7 July 2020

Accepted 20 July 2020

Published 25 August 2020

many healthy women lacking BV symptoms have vaginal microbiomes dominated by microbes associated with BV, resulting in debate about the definition of a healthy vaginal microbiome. Despite substantial evidence that the reproductive health of a woman depends on the vaginal microbiota, future therapies that may improve reproductive health outcomes are stalled due to limited understanding surrounding the ecology of the vaginal microbiome. Here, we use sequencing and metabolomic techniques to show novel associations between vaginal microbes and metabolites during healthy pregnancy. We speculate these associations underlie microbiome dynamics and may contribute to a better understanding of transitions between alternative vaginal microbiome compositions.

KEYWORDS *Lactobacillus*, longitudinal, metabolome, microbiome, pregnancy, vagina

Vaginal microbes sustain important physiologies and produce metabolites that can directly or indirectly affect maternal health and infant development during pregnancy. Perturbations to early-life microbiomes and associated metabolic dysfunction have been linked with allergy and autoimmune diseases such as asthma (1–4). For example, regular prenatal and postnatal farm exposure, i.e., contact with a diversity of microbes during pregnancy and infancy, have been shown to reduce the incidence of chronic health diseases such as asthma and atopy (5). Moreover, recent research has supported the idea of fetal programming, a term describing the process by which the maternal microbiota, as well as maternal antibodies, prepare the infant immune system for the postnatal onslaught of colonizing microbes (6). Others have shown in mice that vaginal dysbiosis, induced by maternal stress, has the potential to negatively affect offspring metabolic profiles (7). Thus, maternal microbes, particularly those of the vaginal tract, are some of the first microbes the offspring will encounter and may be central to the study of early-life microbiome and immune development (8–12). Indeed, a recent large-scale study of 2,582 women, over 600 of whom were pregnant (a subset of whom were longitudinally sampled), provided evidence for vaginal microbiome restructuring during pregnancy toward a *Lactobacillus*-dominated community (13). This occurred early in gestation and was associated with a reduced vaginal microbiome metabolic capacity. Postpartum, irrespective of the mode of delivery, the vaginal microbiota resembled that of a gastrointestinal microbiome, likely due to microbial mixing during the birthing process (9), suggesting that both vaginal and gastrointestinal microbial seeding of the neonate occurs.

The human vaginal microbiome maintains low diversity in low-pH conditions and depends on host sugars as carbon sources, with less access to dietary and exogenous nutrients than the gut, skin, or oral cavity. Historically, vaginal microbial communities have been stratified based on hierarchical clustering of the taxa composition (12). Keystone species include *Lactobacillus crispatus* and *Lactobacillus gasseri*, which have been associated with maintenance of a simple vaginal microbiome by their production of bacteriostatic and bactericidal compounds (e.g., lactic acid and hydrogen peroxide) and maintenance of a low pH (14–16) numerically and functionally dominating their respective vaginal communities. A closely related species, *Lactobacillus iners*, has been associated with health-promoting benefits; however, its genome also encodes the capacity to promote microbiome perturbation by increasing vaginal pH and producing species-specific virulence factors (14, 17–19). Bacterial vaginosis (BV), the most common gynecological condition in reproductive-age women (20), is characterized by the presence of a more diverse vaginal microbiome and associated with adverse pregnancy outcomes, including preterm birth (21), endometritis (22, 23), and spontaneous abortion (24–27). Recently, vaginal microbial transplants have been successfully implemented as a treatment for intractable BV (28). Despite *L. crispatus* generally being regarded as a highly beneficial and dominant microbe throughout pregnancy, healthy women from different ethnic groups have markedly different species dominating the vaginal microbiome (15). In fact, many healthy women who lack BV symptoms have vaginal microbiomes dominated by microbes that are associated with BV (29), suggest-

Time Series

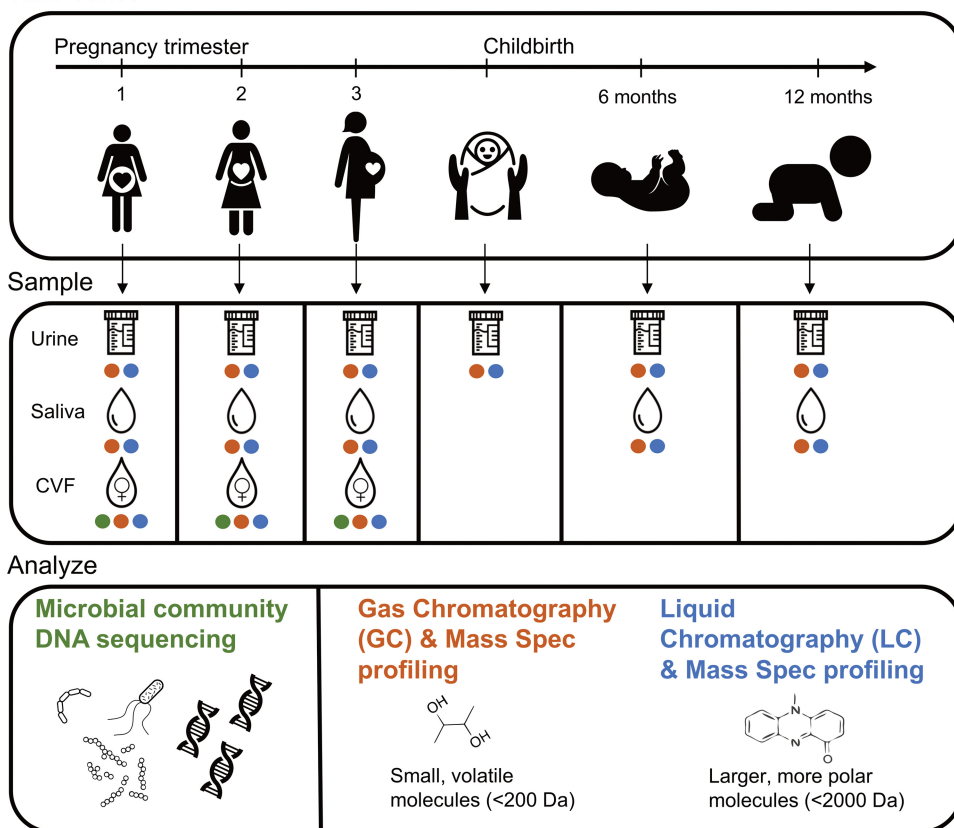


FIG 1 Study outline. Eighteen women were sampled throughout pregnancy and their offspring were sampled at birth and 6 and 12 months of age. Samples collected were urine, saliva, and cervical vaginal fluid (CVF) from the mothers and urine and saliva from the children. CVF was sequenced using shotgun metagenomics and amplicon sequencing. All samples were analyzed using GC-TOF MS and lipidomics.

ing that taxonomy alone is insufficient to predict health outcomes and that microbial activities, including metabolic productivity, may offer a more contemporary view of microbiome function.

An untargeted, more global assessment of microbiomes and associated metabolites during pregnancy and early life is lacking. To address this gap in knowledge, we collected saliva, urine, and cervical vaginal fluid (CVF) from 18 mothers during each trimester of pregnancy and saliva and urine from offspring through their first year of life. Specifically, we were interested in how maternal CVF microbiome profiles are associated with metabolomic assessments of the same samples. Furthermore, we had the opportunity to examine whether maternal saliva and urine metabolome profiles relate to those of the infant in the first year of life. Here, we present DNA sequencing (amplicon and shotgun) and untargeted metabolomics to characterize microbial and metabolic features of the CVF microbiome throughout pregnancy to determine the composition of the vaginal microbiome from a cohort of healthy Caucasian and Hispanic women, longitudinally sampled throughout a healthy pregnancy.

RESULTS

Description of the cohort and data obtained from samples. Saliva, urine, and cervical vaginal fluid (CVF) were collected from 18 women at early, middle, and late pregnancy with the gestational age range of the included women at each time point (Fig. 1). At the time of enrollment into the cohort, the average woman's age was 27.8 years old, and the average prepregnancy body mass index (BMI) was 24.8 kg/m² (Table 1). The cohort was 39% white Hispanic and 61% non-Hispanic white; there were

TABLE 1 Demographics of the 18 mothers who participated in the study

| Maternal ID | Race and ethnicity | Age (yr) | BMI (prepregnancy [kg/m ²]) |
|-------------|--------------------|----------|---|
| 1018 | White Hispanic | 35 | 27.4 |
| 1062 | White Hispanic | 23 | 25.3 |
| 1088 | White non-Hispanic | 26 | 25.8 |
| 1089 | White non-Hispanic | 22 | 21.8 |
| 1103 | White non-Hispanic | 27 | 24.5 |
| 1111 | White Hispanic | 38 | 27.9 |
| 1120 | White non-Hispanic | 34 | 26.9 |
| 1126 | White Hispanic | 19 | 27.8 |
| 1137 | White Hispanic | 31 | 23.5 |
| 1146 | White non-Hispanic | 29 | 23.5 |
| 1151 | White non-Hispanic | 30 | 22.4 |
| 1157 | White non-Hispanic | 28 | 18.9 |
| 1180 | White non-Hispanic | 29 | 24.9 |
| 1191 | White Hispanic | 31 | 22.7 |
| 1198 | White non-Hispanic | 26 | 24.8 |
| 1201 | White non-Hispanic | 30 | 29.9 |
| 1202 | White Hispanic | 20 | 24.7 |
| 1222 | White non-Hispanic | 24 | 24.0 |

no significant differences in BMI (t test; $P = 0.28$) or age (t test; $P = 0.89$) between ethnic groups. Saliva and urine were collected at indicated intervals from each infant up until 1 year of age (Fig. 1). Saliva, urine, and CVF samples were subjected to metabolomics analysis, whereas only CVF was used for sequence analysis. Sequence analysis included amplicon-based sequencing of the 16S rRNA gene (bacteria and archaea) and ITS2 (fungi) loci and shotgun metagenomic sequencing of the entire microbial community.

Vaginal microbiota support high abundances of *Lactobacillus* and *Bifidobacteriaceae* throughout pregnancy. Sixteen individuals (42 total samples) produced sufficient sequence reads for taxonomic assignment using the 16S rRNA gene. Amplicon sequencing stratified cervical samples into those where the most abundant taxon was *Lactobacillus* spp. (34/42; 81%) or *Gardnerella* spp. (8/42; 19%) (Fig. 2A; Fig. S1A). The bacterial taxa in samples with abundant *Lactobacillus* spp. were significantly less evenly distributed (linear mixed-effects [LME] modeling; $P = 0.001$; Fig. 2C), with *Gardnerella vaginalis* being the most abundant in seven of eight samples (88% relative abundance) and a *Shuttleworthia* taxon being most abundant in one sample (at 23% relative abundance). In samples where *Lactobacillus* spp. had the highest abundance, a single taxon comprised 50% or more of the sequencing reads (27/34; 79%). *L. iners* was the most abundant taxon detected in 14 samples from 7 subjects, with a median relative abundance of 79%. In 12 samples from 7 subjects, a *Lactobacillus* taxon, putatively identified as *L. crispatus* through metagenomic sequencing (Fig. S1B and C), had a median relative abundance of 96% and persisted at a relative abundance greater than 90% in subjects 1088, 1120, and 1191. Altogether, the most abundant 3 taxa, *L. iners* (operational taxonomic unit 1 [OTU_1]), *L. crispatus* (OTU_2), and *G. vaginalis* (OTU_3), comprised 66% of the total bacterial sequencing reads.

Whole-genome shotgun sequencing produced, on average, 2.9 million paired-end reads per sample, which decreased to an average of 220,355 paired-end reads per sample following removal of reads that aligned to the human genome. Thirteen individuals (35 total samples) produced sufficient sequence reads for taxonomic assignment, which was concordant with 16S rRNA gene sequence results (Fig. S1B and C). Most of the reads classified as *L. crispatus* or *L. iners* mapped to a single metagenomic assembled genome, with completeness of 95.7% and 97.1% and redundancy of 0% and 1.4%, respectively.

Twelve samples from seven subjects produced ITS2 sequences (Fig. 2A); we do not have quantitative data characterizing the abundance of bacterial or fungal biomass. Eleven samples from six subjects contained species of *Candida*, classified as *C. albicans* (Fig. S2A), the most abundant fungal taxon in these data. Shotgun metagenomics confirmed these results and allowed for additional identification of reads mapping to

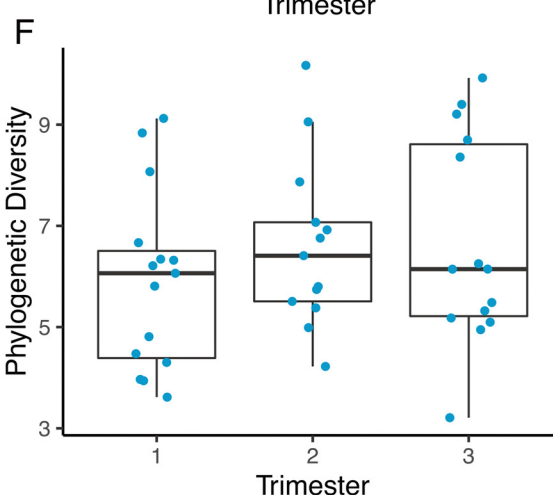
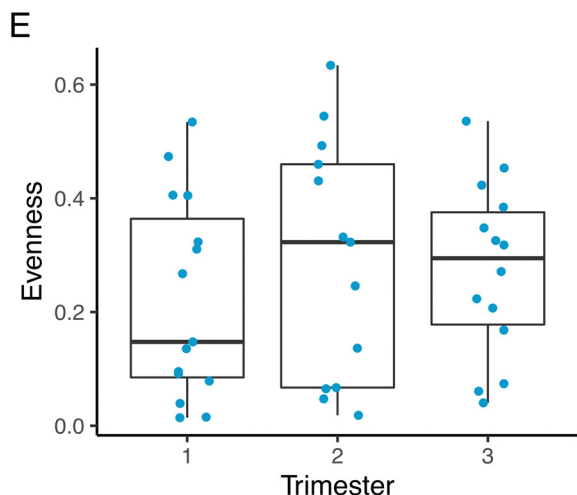
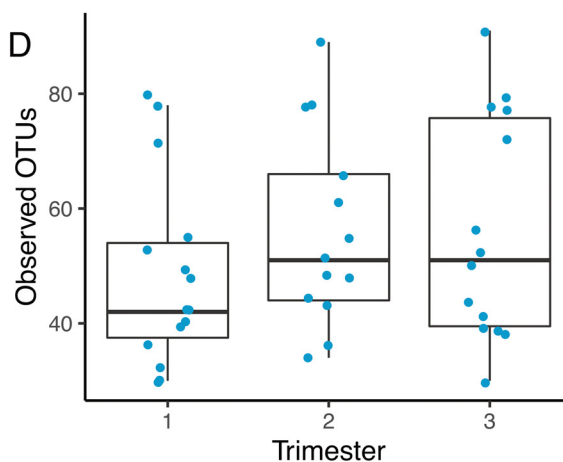
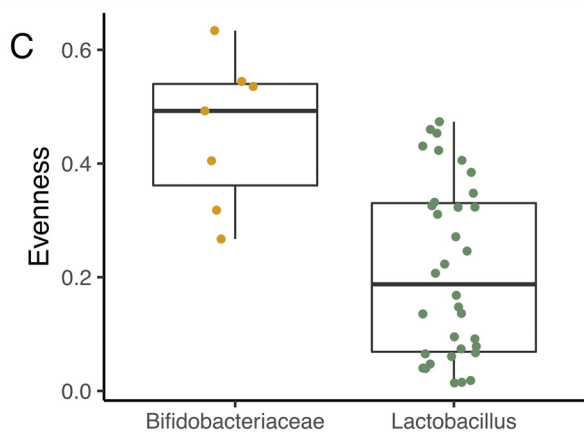
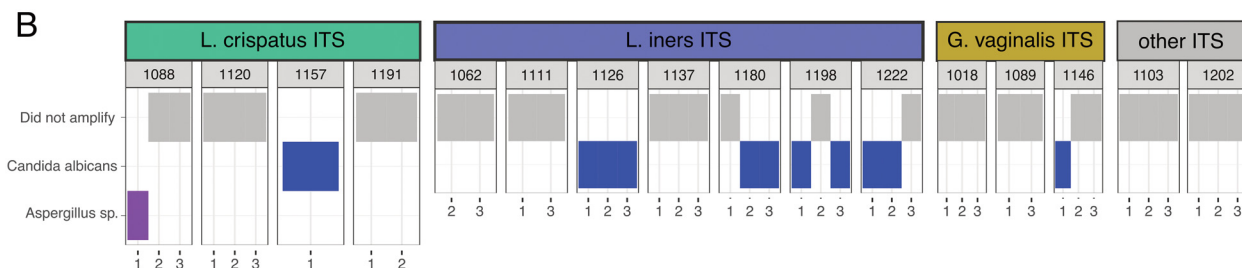
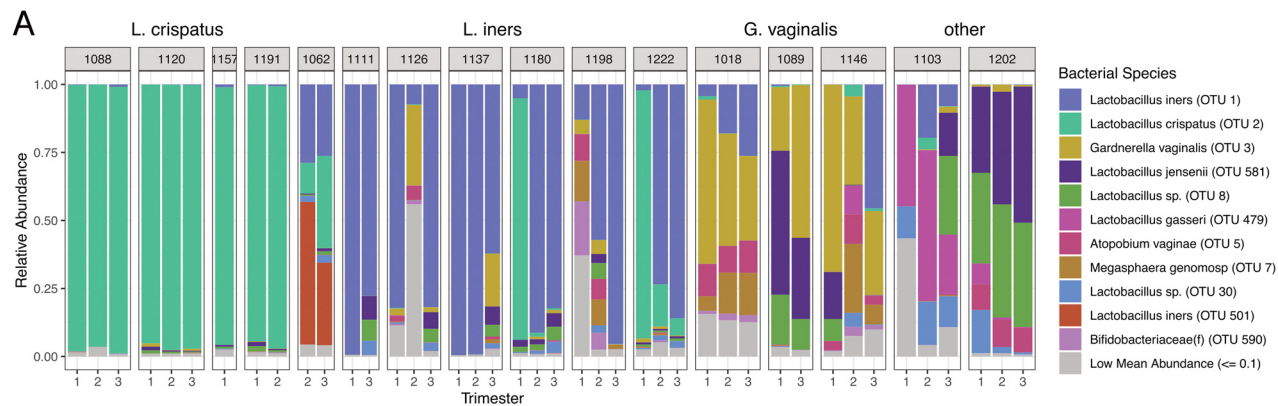


FIG 2 Taxonomy and alpha diversity of vaginal microbiomes during pregnancy. (A) Relative abundance plot of operational taxonomic units, from 16S amplicon data, grouped together by individual. Each individual is clustered into a larger category defined by the dominating microbe. (B) Presence or (Continued on next page)

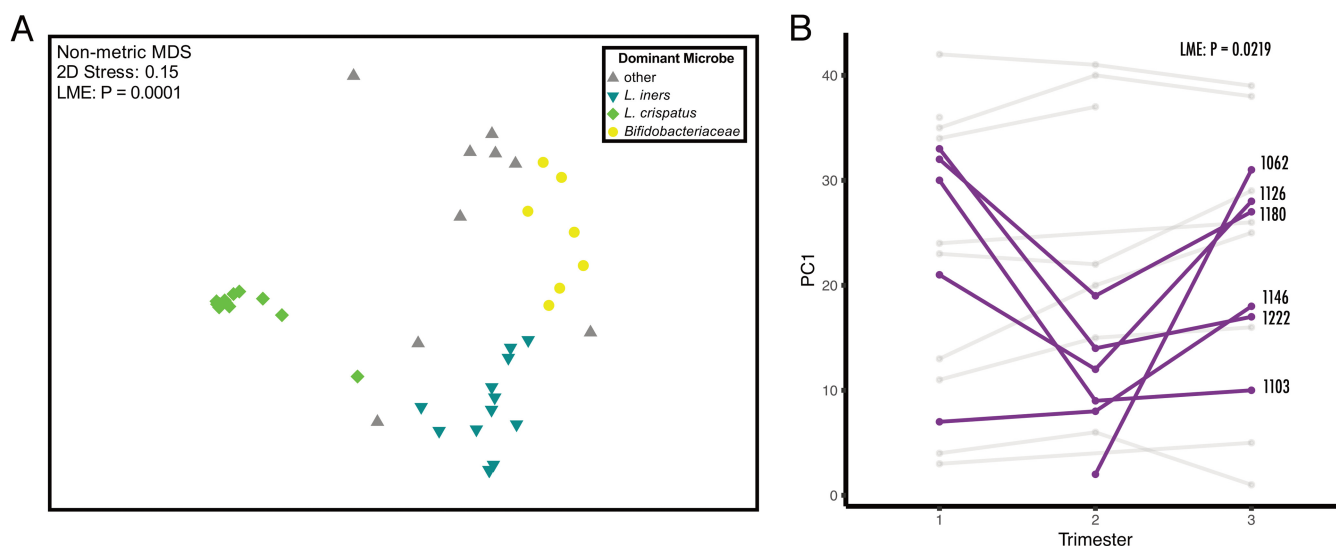


FIG 3 Ordination of vaginal microbiomes during pregnancy. (A) Nonmetric multidimensional scaling (nMDS) of Bray-Curtis dissimilarity between vaginal microbiomes ($n = 42$ samples) of mothers. Color indicates the most abundant microbe within the microbial community. The most abundant microbe in the community plays a statistically significant role in the composition of the community (LME; $R^2 = 69\%$; $P < 0.0001$). (B) Some participants (6/16 individuals) experienced large, significant shifts (LME; $P = 0.0219$) in their microbiomes throughout the trimesters of pregnancy.

taxa such as *Malassezia* spp. (Fig. S2C). Subject 1,088 was the only participant to deviate from this trend, with a high relative abundance of *Aspergillus* during the first trimester of pregnancy (Fig. S2A).

Alpha diversity indices based on 16S rRNA and ITS2 data, when available, were compared across trimesters. While some subjects exhibited qualitative evidence of compositional shifts in vaginal microbiota with advancing gestation (Fig. 2A), we did not observe a significant difference in bacterial richness (number of observed OTUs) (LME; $P = 0.17$), evenness (Pielou's evenness index) (30) (LME; $P = 0.46$), or phylogenetic diversity (LME; $P = 0.21$) across trimesters (Fig. 2D to F).

Highly abundant bacterial taxa were significantly associated with community composition. A nonmetric multidimensional scaling (nMDS) plot of Bray-Curtis dissimilarities showed that vaginal communities clustered by their most abundant bacterium (Fig. 3A). This association between the most abundant bacterial taxa and microbial composition of the sample was significant and explained more than half of the variance using permutational multivariate analysis of variance (PERMANOVA) ($R^2 = 56\%$; $P = 0.0001$). To account for repeated measures from longitudinal samples from the same individual, we also performed an LME, which required dimensional reduction (LME; $R^2 = 69\%$; $P < 0.0001$). Communities with abundant *L. crispatus* were more similar to each other, sharing more than 90% similarity, in comparison to communities where a different bacterial species was most abundant. While some individuals exhibited a relatively stable microbial community over time, others (6/16) experienced shifts in composition, resulting in a statistically significant change in Bray-Curtis dissimilarity on PC axis 1 between trimesters (Fig. 3B).

The five subjects with *Candida* detected in at least one of two longitudinally paired samples displayed a significant increase in intersample Bray-Curtis dissimilarity in their bacterial profiles over that interval (e.g., the intervals between trimester 1 to 2, 2 to 3, or 1 to 3), suggesting the presence of *Candida* may be associated with greater shifts in bacterial composition than those who had no *Candida* detected (Fig. S2B).

FIG 2 Legend (Continued)

absence of fungi, at the genus level, per sample. Linear mixed-effects models (LME) were done on the alpha diversity metrics to account for repeated measures in the data. (C) Evenness between samples dominated by *Lactobacillus* ($n = 34$ samples) is significantly lower than samples dominated by *Bifidobacteriaceae* ($n = 7$ samples). (D) No significant change in the observed OTUs between the trimesters ($n = 15, 13,$ and 14 samples, respectively) of pregnancy and likewise. (E and F) There was no change in evenness (E) or phylogenetic diversity (F) throughout pregnancy.

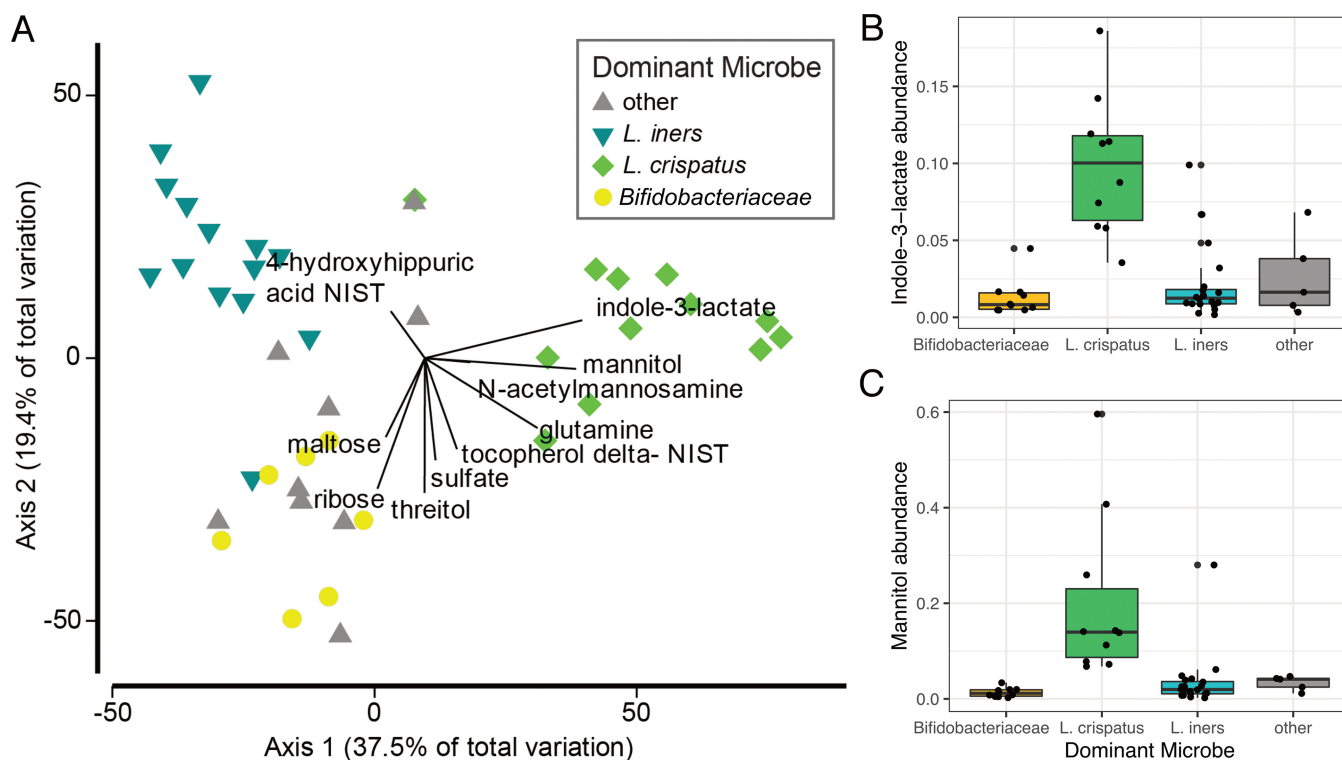


FIG 4 Relationship between vaginal microbes and metabolites. (A) Distance-based linear model recapitulates the relationship between the vaginal microbiomes of these subjects ($n = 42$ samples). Superimposed are vectors showing which annotated GC-TOF MS molecules are best correlated with these microbial communities. Length and direction of vectors correspond to the strength of the association between the metabolite and the microbial communities. Box plots show the raw abundance ($n = 45$ samples) of indole-3-lactate (B) and mannitol (C).

Metabolites have strong associations with vaginal microbial community structures.

Using gas chromatography-time of flight mass spectrometry (GC-TOF MS), we detected 330 metabolites from urine, saliva, and CVF with 133 identified compounds. In the same samples, 1,946 metabolites were also detected by liquid chromatography-quadrupole time of flight tandem mass spectrometry (LC-QTOF MS/MS) (lipidomics; Table S1), with an additional 353 identified compounds. The CVF metabolome as assessed by both mass spectrometry methods did not significantly differ across trimesters (LME; $P = 0.6378$ for GC-TOF MS; $P = 0.3942$ for liquid chromatography-tandem mass spectrometry [LC-MS/MS]). This stability was even true for the subset of individuals who exhibited shifts in microbiota composition over trimesters (LME; $P = 0.6594$ for GC-TOF MS; $P = 0.2482$ for LC-MS/MS). CVF samples dominated by distinct bacteria exhibited significant differences in metabolic profiles (PERMANOVA; $R^2 = 12\%$; $P = 0.0195$). A constrained, distance-based ordination plot recapitulated 67% of the community variation observed in the vaginal microbiota (Fig. 4). Superimposed on the ordination plot are GC-TOF MS predictor metabolites, calculated using the distance-based linear models (DISTLM) program in PRIMER-e. Indole-3-lactate (ILA) accounted for 27% of the variation observed in the vaginal microbiota data and was found to be more abundant in vaginal microbiota with abundant *L. crispatus* (Fig. 4B). Mannitol was also more abundant in samples dominated by *L. crispatus* (Fig. 4C). In parallel, we found that a pathway for mannitol-1-phosphate production is also more abundant in shotgun metagenomic data sets of CVF samples dominated by *L. crispatus* (Fig. S3). This linear model identified the top 10 annotated GC-TOF MS metabolites that were associated with variation in the microbial community composition, which are shown in Fig. 4A; these 10 metabolites together might explain almost 57% of the total variation in the microbial community composition. A permuted random forest recapitulated what we found in the DISTLM, identifying mannitol and indole-3-lactate as two of the top variables of importance, specifically for distinguishing microbiomes with high abun-

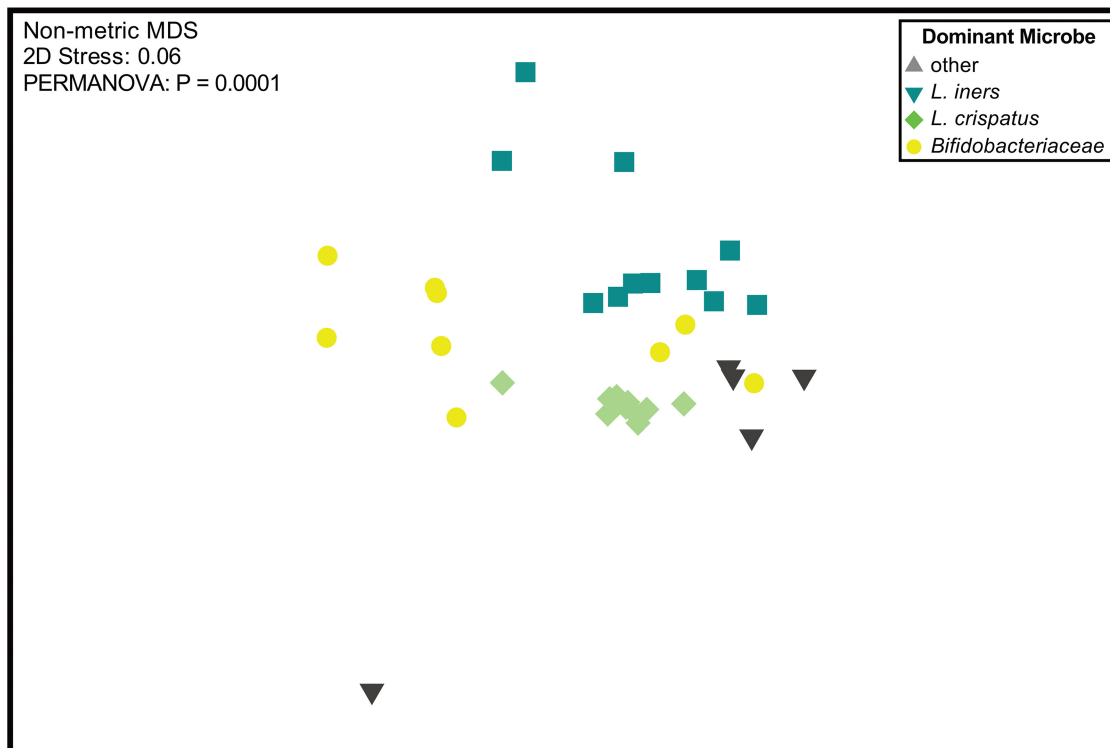


FIG 5 Ordination of functional pathways within the vaginal microbiome. An nMDS of HUMAnN2 analysis, examining the abundance of pathways in each microbiome ($n = 35$ samples). Vaginal microbiomes have functions that are indicative of the most abundant microbe present in the samples.

dances of *L. crispatus* (Fig. S4A). To explore the ability to analyze the metabolome in high throughput, the same sample sets were analyzed using matrix-assisted laser desorption ionization mass spectrometry imaging (MALDI-MSI) (Table S1). Detected ions by MALDI were compared to those identified by GC-MS and LC-MS, and we found that ~55% of the metabolites identified had corresponding ions in the MALDI analysis (Table S1).

Metagenomics and functional potential of communities. Distinct functions were associated with each of the vaginal microbial community clusters (PERMANOVA; $R^2 = 70\%$; $P = 0.0001$; Fig. 5). LEfSe identified several pathways that differed between *L. crispatus* and *G. vaginalis*, in particular, an enrichment of ammonia assimilation genes in *G. vaginalis* (Fig. S4B). Genes involved in mannitol metabolism were enriched in communities where *L. crispatus* was highly abundant (Fig. S3C and Fig. 4B). Searching the PATRIC database of all sequenced *L. crispatus* (64 genomes), *L. iners* (22 genomes), and *G. vaginalis* (127 genomes) strains revealed annotated genes for mannitol usage and transport for *L. crispatus* but not for *L. iners* or *G. vaginalis*.

Mothers and infants have significantly different saliva and urine metabolomes. Maternal and infant saliva and urine metabolomes were assessed with both GC-TOF MS and LC-MS/MS (lipidomics) in order to study the relationship between maternal and infant metabolomic compositions during early life (see Fig. 1). PERMANOVA analysis of lipidomics data from saliva samples showed the largest difference between mothers and offspring (PERMANOVA; $R^2 = 69\%$; $P = 0.001$; Fig. S5A). A subset of 50 lipidomics metabolites with high mean abundance, 70% of which were unannotated, showed distinct profiles between mother and offspring salivary metabolomes (Fig. S6A). Likewise, GC-TOF MS salivary metabolomes were also significantly different between mother and offspring, but far less variation was explained (PERMANOVA; $R^2 = 12\%$; $P = 0.0001$; Fig. S5B). Some metabolites, such as lactulose, were much more abundant in infants and largely absent in mothers (Fig. S6B). Maternal metabolomics profiles

(both GC-TOF MS and LC-MS/MS lipidomics) have a strong individual signature, while infants do not (see PERMANOVAs; Table S2). The infant metabolome for saliva and urine had little variance attributed to which subject donated the sample, but GC-TOF MS was able to detect a significant change between the infant urine metabolome at birth versus 6 and 12 months of age (PERMANOVA; $R^2 = 34\%$; $P = 0.0007$; Table S2). Moreover, from lipidomics data, the infant metabolome profile seemed to converge on mothers' metabolomes as they aged, though more samples would be needed to confirm this finding (Fig. S5C). For both saliva and urine, GC-TOF MS- and lipidomics-detected metabolites were more similar for mother-child pairs than for unrelated individuals (Fig. S7). Mantel tests to determine if intersample relationships were similar between chromatography methods (including both GC-TOF MS and lipidomics) showed a strong correlation between saliva samples and weaker correlations between urine and CVF (Table S3).

DISCUSSION

Exposure to the microbiome in early life is critical for immune and physiological development (1–4), yet the factors that set this trajectory remain poorly understood. In this study, we followed the vaginal microbiome through the trimesters of pregnancy for 18 women, tracking changes in the bacterial communities with longitudinal samples and capturing their functional potential with metagenomic sequencing and multiple platforms to assess metabolomic profiles. The resolution provided by shotgun metagenomic sequencing allowed us to identify species and characterize the functional gene content of CVF microbiomes. An additional strength of this work is the strict inclusion criteria defining healthy pregnancy (see “Subject information” in Materials and Methods). Moreover, as part of an existing sample cohort, we had the opportunity to measure saliva and urine metabolomes from mothers and children. We aim to establish how the metabolome develops in the first year of life and how maternal-infant saliva and urine metabolomes relate. In our study, most healthy pregnant women exhibited a relatively stable vaginal microbiota throughout the trimesters of pregnancy, dominated by *Lactobacillus* or, in some cases, more diverse, *Bifidobacterium*-dominated microbiota. However, a subset of women exhibited compositional shifts in their CVF microbiota as pregnancy progressed, as has also been seen in other larger cohorts (13). We found several strong correlations between particular vaginal communities and metabolites, which may help us understand the physiology underlying distinct vaginal microbiota structures that were evident in our study. Lastly, vaginal microbiota composition predicted which metabolites were present in the CVF samples but not urine or saliva samples from the mothers or infants, suggesting that local microbial metabolism may represent the dominant contributor to the metabolic milieu of the vaginal tract during pregnancy.

Our study supports the results from several other studies that have indicated that the vaginal microbiome is stable during pregnancy (8, 31). Specifically, in a longitudinal study that includes 90 women, most retain a microbial community with the same dominant member (in *L. crispatus* communities, 75% remain stable; in communities with high *L. iners* abundance, 71% remain stable; and in more diverse communities like those sometimes associated with BV, 58% do not shift) (13). Using metagenomic sequencing to probe microbial community variation, our findings indicate that few bacteria, particularly *Lactobacillus* species, are highly abundant in the vaginal environment. Indeed, for individuals with vaginal microbiomes numerically dominated by *L. crispatus* or *L. iners*, the vast majority of reads mapped to contigs from one strain of *L. crispatus* or *L. iners* (Fig. S1B). Of note, the microbiome of some individuals did differ considerably with advancing pregnancy. For instance, the vaginal microbiota of subjects 1180 and 1222 had higher abundances of *L. crispatus* during the first trimester, but *L. iners* was more abundant in the remaining trimesters. Brooks et al. (32) demonstrated that shifts in vaginal microbiota structures can be described probabilistically, where shifts from *L. crispatus* to *L. iners* are the most likely to occur. This is consistent with the observations made in two individuals from our study; however, due to the small

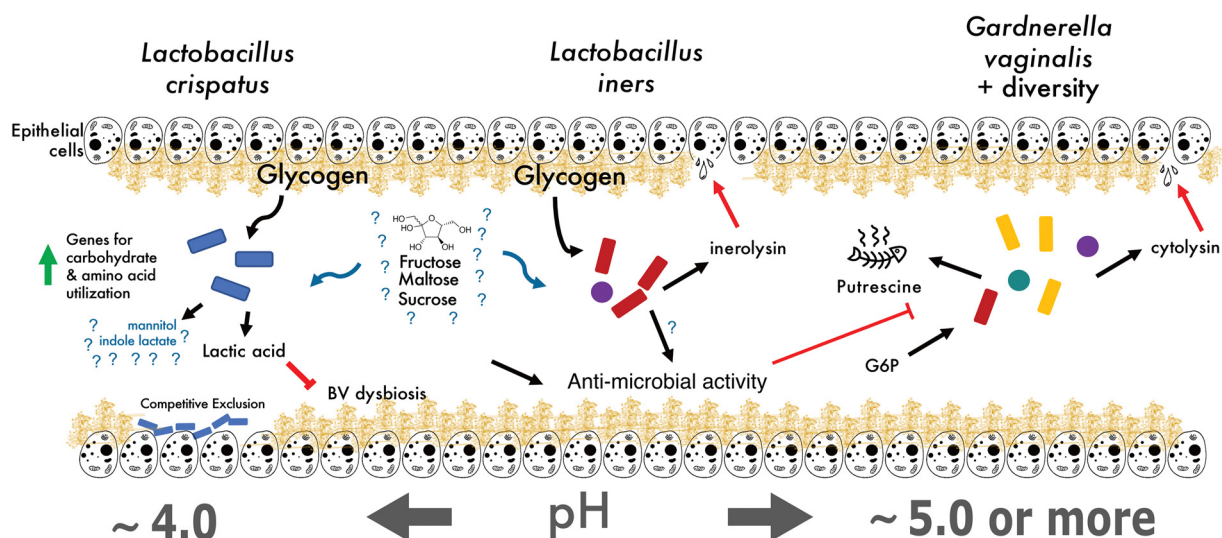


FIG 6 Proposed vaginal microbial community model. Current hypothesized model of vaginal microbial community physiology, with gaps in understanding (denoted by question marks) where future work is needed. Our study indicates that mannitol production is associated with a high relative abundance of *L. crispatus*.

number of samples exhibiting this phenomenon, qualitative assessments were more appropriate than statistical analysis. Of note, vaginal microbiota instability throughout pregnancy was associated with the presence of *Candida*, a known opportunistic pathogen of the vaginal tract. Since inclusion criteria for this study stipulated no antibiotic treatment, it is unlikely that *Candida* detection was a result of antibiotic administration. The prevalence of *Candida* in our cohort is more likely to be reflective of the fact that pregnancy is a known risk factor for candidiasis (33) and to related differences in the vaginal environment, including microbiological colonization. Indeed, *L. crispatus* has been shown to have anti-*Candida* activity (34), and 90% (10/11) of samples that were *Candida* positive came from individuals whose vaginal microbiota were dominated by an organism other than *L. crispatus*.

A few metabolites were highly indicative of the bacterial community present in each subject and may be useful biomarkers for the type of vaginal microbiota present. The most indicative metabolite was indole-3-lactate (ILA), a tryptophan metabolite whose abundance was correlated with communities having abundant *L. crispatus* (Fig. 4B and Fig. S4A). One potential explanation is that *L. crispatus* produces ILA to competitively exclude the growth of other species (Fig. 6). At physiologically relevant concentrations, ILA has been shown to have antimicrobial properties against both Gram-positive and Gram-negative organisms (35, 36). Although the production of lactic acid is generally thought of as a strategy *Lactobacillus* spp. use to prevent other species from colonizing the vagina, perhaps these organisms also use ILA in a similar or supplementary capacity. Additionally, bacteria-derived ILA (also referred to as indole-lactic acid) has been recently shown to directly move from maternal to infant tissue (6). It has been suggested that indoles may play an important role as a ligand for the human aryl hydrocarbon receptor (AhR), which have diverse functions from immune regulation to metabolism (reviewed in reference 37). Zelante et al. further showed that some lactobacilli produce the related tryptophan catabolite, indole-3-aldehyde (IAld), which provides protection against candidiasis by increasing interleukin-22 (IL-22) production via AhR receptor binding (38). The study also demonstrated that vaginal specific bacteria, such as *L. acidophilus*, produce IAld in the vaginal environment, which protected against vaginal but not intestinal candidiasis. We measured indole-3-acetate (IAA), the direct precursor to IAld, in our study, but found no difference in its abundance between women dominated by different species of *Lactobacillus* (data not shown). Because indole-3-lactate can act as a ligand for AhR, we speculate that *L. crispatus* may

regulate the IL-22-AhR response in the vagina, reducing the risk of vaginal candidiasis in the same way IALd does, and potentially activating the AhR response in newborns to prevent early-life candidiasis. Additionally, indole itself may play a role in structuring community composition by selecting for organisms that have adapted to a high abundance of this metabolite, repelling more transient microbes that have not been exposed to higher indole concentrations previously (39).

Whole-genome shotgun metagenomics allowed us to begin to address the functional capacities of these microbiomes. The largest differences between functional capacity appeared to be between communities where *L. crispatus* or *G. vaginalis* were the most abundant bacterial taxon. One pathway particularly enriched in *Gardnerella* communities was the ammonia assimilation cycle (Fig. S4B). Studies have pointed out *Gardnerella's* preference for ammonia as a nitrogen source (40); moreover, this ability to assimilate ammonia and produce amino acids has been implicated in mutualistic interactions between species of *Prevotella*, in the context of BV (41). Together, these BV-associated organisms contribute to genital inflammation, which may play a role in the susceptibility of certain diseases, such as HIV (42).

Increased abundance of mannitol when *L. crispatus* was present is an important and unexpected finding (Fig. 4). Most likely, mannitol contributes to optimizing the tonicity of the vaginal environment, and has recently been considered for this use in developing effective therapeutics for altering the vaginal microbiota (43, 44). Even more, mannitol may assist *L. crispatus* in adhering to the epithelial layer, a strategy the organism may use to competitively inhibit other microbes from colonizing, potentially by drawing out excess water in the mucin layer and altering the mucin structure (45). Irrespective of the biochemistry, these genes, and mannitol in general, represent very specific markers of a community where *L. crispatus* was most abundant.

Although it is known that homofermentative lactic acid bacteria (LAB) such as *L. crispatus* (46) convert glucose primarily to lactic acid, it is unclear why mannitol accumulates in this niche. Interestingly, there was no difference in the glucose abundance between the four distinct vaginal communities. Further, metabolomic analysis of our CVF samples failed to capture significant levels of the mannitol precursor fructose; however, previous studies have indicated an appreciable amount of fructose within the cervical mucus of humans (47) and the capability of *L. crispatus* to utilize fructose as a carbon source (48, 49). We speculate that this high extracellular mannitol abundance phenotype may underlie the cell's need to regenerate NAD⁺ for use in glycolysis. When faced with a limiting amount of pyruvate (or perhaps an upstream glycolytic metabolite) to convert to lactate, homofermentative LAB may be unable to produce sufficient NAD⁺ to allow glycolysis to continue. To this end, reducing fructose 6-phosphate to mannitol-1-phosphate may be an alternative and vital way *L. crispatus* regenerates NAD⁺ for glycolysis (50, 51). We did find the gene mannitol-1-phosphate dehydrogenase, responsible for converting fructose 6-phosphate to mannitol-1-phosphate, was highly correlated ($R^2 = 0.9$) with the relative abundance of *L. crispatus*. The genes for the conversion of mannitol-1-phosphate to mannitol (presumably an M1P phosphatase) and its subsequent export are currently unknown (51). This may imply that mannitol accumulation is a marker of a cellular switch to NAD⁺ regeneration by fructose reduction rather than converting pyruvate to lactic acid. Consequently, the decrease in lactic acid production may contribute to community dysbiosis due to a rise in pH (Fig. 6). Future experiments using culturing to elucidate whether these *in vivo* community data are recapitulated with axenic cultures *in vitro* are needed.

Furthermore, this study enabled comparison of two metabolomic methods (GC-TOF MS and LC-MS/MS lipidomics) for analysis of the pregnancy and early-life metabolomes. Our data showed that metabolite intensities obtained by GC-TOF MS were more tightly correlated with microbial community composition than those obtained by lipidomics, perhaps indicating that GC-TOF MS is more effective at detecting microbial metabolites than lipidomics, especially during pregnancy. We also show that both the saliva and urine metabolomes of children are more similar to their own mother than to unrelated individuals (Fig. 6). Strikingly, the ability of lipidomics to differentiate mothers from

children via saliva was the strongest signal in our metabolomics data (Table S1). The oral microbiome may play a role in this, as there is a well-established community succession in children during early life (reviewed in reference 52), where children begin life with oral microbes that differ from those in adults. Lactulose, detected by GC-TOF MS, was a specific metabolite with increased abundance in infant saliva, which may reflect its use as a treatment for constipation (53) or perhaps even its presence in heated milk (54). Finally, urine metabolomes had a distinct age profile, especially with the lipidomics data. Our data suggest that, over the first year of life, the urine metabolome rapidly converges on the adult metabolome. This is likely a result of the development of the renal system in children (55), along with the development of the gut microbiome and the related metabolites which are processed through the liver and kidneys. Other reasons for this age-related shift include a change in diet and weaning off breastmilk or formula (55). Expectedly, we did not see a strong influence of the vaginal microbiome during pregnancy on the infant saliva and urine metabolome. We suspect that if differences in the vaginal microbiome were to affect the early-life saliva and urine metabolome, those effects would be subtle. The lack of stool samples from the mothers and infants is a limitation of this study, as they may contain a stronger signal of shared metabolomes across mother-infant dyads. Additionally, we explored a high-throughput approach to characterize the metabolomes. By using acoustic deposition in combination with MALDI-MSI, a throughput of ~ 1 s per sample was reached using only $2 \mu\text{l}$ of sample. Of the metabolites identified by GC-MS and LC-MS, $\sim 55\%$ had corresponding ions in the MALDI-MSI analysis (Table S1). Future work will focus on confirming these metabolite identifications, but the initial results are promising and indicate that rapid analysis of microbial metabolites using MALDI, an analysis platform routinely used in clinical microbiology laboratories, is feasible (56).

Overall, we provide a broad look at the metabolome during pregnancy and early life, detailing the utility of GC-TOF MS, lipidomics, and MALDI-MSI for saliva, urine, and CVF.

In conclusion, here, we share a high-resolution characterization of the vaginal microbiome, longitudinally sampled throughout healthy pregnancies. We show that, despite the generally accepted view that lactobacilli are indicative of healthy vaginal communities, many women in our healthy pregnancy cohort had non-*Lactobacillus*-dominated communities. The vaginal communities were characterized by a high abundance of one or a few acid-tolerant species, which dictated the physiologic potential and the metabolic profiles of the vaginal microbiome. Many of the metabolites that were specific to these different organisms warrant further investigation, especially considering the recent development of vaginal microbiota transplantation (VMT) as a treatment for BV (28). The metabolites we found to be associated with *L. crispatus* may be useful as microbiome cultivation approaches are developed to intentionally direct the composition of the vaginal microbiome. For example, indole-3-lactate may support *L. crispatus* colonization, while mannitol may indicate a shift in metabolism away from fermentation and the production of acid, relaxing the low-pH selection pressure, which normally gives *L. crispatus* an advantage.

MATERIALS AND METHODS

Subject information. Eighteen women were selected from a larger cohort recruited to address how maternal stress affects child development (57, 58) (Table 1). Inclusion criteria for the larger cohort included >18 years of age, singleton, intrauterine pregnancy, and nondiabetic. Additional inclusion criteria for this study were normal prepregnancy body mass index, vaginal delivery, full-term pregnancy, breastfeeding, and no antibiotics for mother or baby. Generally, these 18 women and their children represented healthy subjects with the most complete sample sets.

Sample collection. Samples were collected at each trimester of pregnancy for women and throughout the first year of life for infants. At each time point, maternal saliva, urine, and cervical vaginal fluid were collected. For infants, urine was collected at birth, 6 months, and 12 months, whereas saliva was sampled at 6 months and 12 months of age. Maternal saliva was collected using a Salivette collection kit, including a small cotton roll contained in a plastic container (Salimetrics, Carlsbad, CA). Mothers were instructed to place the cotton rolls in their mouths until saturated with saliva (approximately 1 to 3 min) and then reseal the swabs in plastic Salivette tubes. Infant saliva was collected using Weck-Cel spears and a swab extraction tube system. Infants were allowed to suck on the spear for 2 min, ensuring saturation. Maternal urine was collected using a sterile collection cup. Infant urine was collected using an adhesive

U-Bag attached to the genital region of the infant. A minimum of 2 ml of urine was collected. Cervical vaginal fluid (CVF) was collected by placing three Dacron swabs into the cervix for 10 s to achieve saturation. Each swab was then placed in a plastic vial with 500 μ l of sterile phosphate-buffered saline (PBS). All samples were initially stored at -20°C , and then saliva, CVF, and infant urine were subsequently moved to -80°C storage.

Metabolomics. Prior to processing, samples were thawed from -80°C storage. Fifty microliters of each sample were subjected to gas chromatography-time of flight mass spectrometry (GC-TOF MS) (59) and liquid chromatography-tandem mass spectrometry (LC-MS/MS, lipidomics). Urine, saliva, and CVF from each time point were sent to the West Coast Metabolomics Center (WCMC) for untargeted metabolomics. GC-TOF MS metabolites were extracted with a mixture of 3:3:2 acetonitrile-isopropyl alcohol-water according to standard operating procedures from the Fiehn Lab at the WCMC (60). LC-MS/MS samples were extracted using a variant of the Matyash method (61). Data were acquired for complex lipids in positive and negative electrospray mode on a Waters CSH column and an Agilent 6530 QTOF mass spectrometer (61). Metabolites were identified by retention time MS/MS matching using MassBank of North America (<http://massbank.us>) and NIST 17 libraries.

High-throughput metabolomics. All urine, saliva, and CVF were analyzed using matrix-assisted laser desorption ionization mass spectrometry imaging (MALDI-MSI) for high-throughput untargeted metabolomics. Extracted samples in 3:3:2 acetonitrile-isopropyl alcohol-water were diluted 1:2 in water in 384-well plates. Next, an equal volume of MALDI matrix (20 mg/ml of 1:1 2,5-dihydroxybenzoic acid and α -cyano-4-hydroxycinnamic acid in 1:3 [vol/vol] H_2O -MeOH plus 0.2% formic acid) was added. Samples were printed onto a stainless steel blank MALDI plate using an ATS-100 acoustic transfer system (BioSera) with a sample deposition volume of 10 nl. Samples were printed in clusters of four replicates, with the microarray spot pitch (center-to-center distance) set at 900 μm . MS-based imaging was performed using an ABI/Sciex 5800 MALDI TOF/TOF mass spectrometer with a laser intensity of 3,500 (arbitrary units) over a mass range of 50 to 3,000 Da. Each position accumulated 20 laser shots. The instrument was controlled using the MALDI-MSI 4800 imaging tool. Surface rasterization was oversampled using a 75- μm step size. Average ion intensity for all reported ions was determined using the OpenMSI Arrayed Analysis Toolkit (OMAT) software package (62).

DNA extraction. Cervical brushes were resuspended in PBS. Two negative extraction controls using sterile PBS were prepared alongside the samples. Aliquots of 100 to 200 μ l were added to lysing matrix E tubes prealiquoted with 500 of hexadecyltrimethylammonium bromide (CTAB) DNA extraction buffer and incubated at 65°C for 15 min. An equal volume of phenol-chloroform-isoamyl alcohol (25:24:1) was added to each tube, and samples were homogenized in a Fast Prep-24 homogenizer at 5.5 m/s for 30 s. Tubes were centrifuged for 5 min at $16,000 \times g$, and the aqueous phase was transferred to individual wells of a 2-ml 96-well plate. An additional 500 μ l of CTAB buffer was added to the lysing matrix E tubes, the previous steps were repeated, and the aqueous phases from paired extractions were combined. An equal volume of chloroform was mixed with each sample, followed by centrifugation at $3,000 \times g$ for 10 min. The aqueous phase (600 μ l) was transferred to a clean 2-ml 96-well plate, combined with 2 volume equivalents of polyethylene glycol (PEG) and stored overnight at 4°C to precipitate DNA. Plates were centrifuged for 60 min at $3,000 \times g$. DNA pellets were washed twice with 300 μ l of 70% ethanol, air-dried for 10 min, and resuspended in 100 μ l of sterile water. DNA was quantified using the Qubit dsDNA high-sensitivity (HS) assay kit and diluted to 10 ng/ μ l when possible. Although DNA was extracted from CVF, attempts to extract DNA from saliva were unsuccessful, potentially due to the storage swabs trapping the biomaterial.

Amplicon gene sequencing. To amplify the V4 region of the bacterial 16S rRNA gene, 10 ng of DNA template was combined with PCR master mix (0.2 mM deoxynucleoside triphosphate [dNTP] mix, 0.56 mg/ml bovine serum albumin [BSA], 0.4 μM Illumina adapter sequence-tagged forward primer [515F] [63], and 0.025 U/ μ l *Taq* DNA polymerase) and 0.4- μM barcode-tagged reverse primers (806R) and then amplified in triplicate 25- μ l reactions, along with a no-template control, for 30 cycles (98°C for 2 min; 98°C for 20 sec, 50°C for 30 sec, and 72°C for 45 sec; repeat steps 2 to 4 29 times; 72°C for 10 min). PCR conditions were identical for ITS2 amplification (primer pair ITS7 [5'-GTGARTCATCGAATCTTTG-3'] and ITS4 [5'-TCTCCGCTATTGATATGC-3']) except for the annealing temperature, which was 52°C . Triplicate reactions were combined and purified using the SequelPrep normalization plate kit (Invitrogen) according to the manufacturer's specifications. Purified amplicons were quantified using the Qubit double-stranded DNA (dsDNA) HS assay kit and pooled at equimolar concentrations. The amplicon library was concentrated using the Agencourt AMPure XP system (Beckman-Coulter), quantified using the Kapa library quantification kit (Kapa Biosystems), and diluted to 2 nM. Equimolar PhiX was added at 40% final volume to the amplicon library; the 16S rRNA amplicon pool was sequenced on the Illumina NextSeq 500 platform on a 153-bp by 153-bp sequencing run, and the ITS2 amplicon pool was sequenced on the Illumina MiSeq platform on a 290-bp by 290-bp run.

Shotgun metagenomics sequencing. Sequencing libraries were prepared using the Illumina Nextera kit and methods described in Baym et al. (64). Briefly, DNA from each sample was diluted to 0.5 ng/ μ l and tagged with the Nextera enzyme (Illumina) for 10 min at 55°C . Following tagmentation, each sample received 1- μ l forward and 1- μ l reverse barcodes, which were added via PCR using Phusion DNA polymerase (New England BioLabs). After PCR, the libraries were cleaned of smaller DNA fragments, using AMPure XP magnetic beads (Beckman-Coulter), and pooled by concentration. Libraries were quantified using the Quanti-iT PicoGreen dsDNA kit (Thermo Fisher Scientific), and DNA was run on a gel to check fragment size. These libraries were loaded onto the Illumina Next-Seq 500 at 1.8-pM concentrations and Illumina's midoutput kit for 75-bp paired-end sequencing.

OTU table generation. Raw sequence data were converted from bcl to fastq format using bcl2fastq v2.16.0.10. Paired sequencing reads with a minimum overlap of 25 bp were merged using FLASH v1.2.11. Index sequences were extracted from successfully merged reads and demultiplexed in the absence of quality filtering in QIIME (Quantitative Insights into Microbial Ecology, v1.9.1), and reads with more than two expected errors were removed using USEARCH's fastq filter (v7.0.1001). Remaining reads were dereplicated, clustered into operational taxonomic units (OTUs) at 97% sequence identity, filtered to remove chimeric sequences, and mapped back to OTUs using USEARCH v8.0.1623. Taxonomy was assigned with the most current Greengenes database for bacteria (63) (May 2013) and UNITE version 6 for fungi (65). OTUs detected in negative extraction controls (NECs) were considered potential contaminants and filtered by subtracting the maximum NEC read count from all samples; any remaining OTU with a total read count less than 0.001% of the total read count across all samples was removed. Sequencing reads were rarefied to an even depth (28,972 reads for 16S; 91,232 reads for ITS2). To maximize similarity between the raw and rarefied OTU tables, random subsampling was performed at predefined depths 100 times, and the most representative subsampled community was selected based on the minimum Euclidean distance to the other OTU vectors generated for each sample.

16S rRNA gene analysis. Alpha diversity indices and Bray-Curtis dissimilarity matrices were generated in QIIME (66). Linear outcomes were assessed by linear mixed-effects (LME) modeling to adjust for repeated measures using the nlme package (67) in the R environment (68). Variables of $P < 0.05$ were considered statistically significant. Data were visualized using Tableau and Adobe Illustrator unless otherwise noted.

Metagenomic analysis. Raw sequences (mean 2,977,881 paired-end reads per sample from 35/38 successfully sequenced samples) were filtered using PRINSEQ v0.20.4 (69) to filter out sequences that had a mean quality score of 30 or less. Human DNA was next filtered out by aligning the filtered reads to the human genome (hg38) using Bowtie2 v2.2.7 (70) and keeping the reads that failed to align (mean 220,355 paired-end plus 33,744 singleton reads per sample or 10.9% of quality-filtered reads per sample). To analyze functional potential, the reads were run through HUMAnN2 v0.1.9 (71) using default parameters, and differences in pathway abundances were analyzed using LEfSe (72). These reads were also cross-assembled using SPAdes v3.8.2 (73). Each sample was then mapped to this cross-assembly using Bowtie2, samples from the same subject were merged together using Samtools v1.9, and the resulting bam files and the cross-assembly were imported into Anvio4 (74). Taxonomy was assigned to each gene call using Kaiju (75), which subsequently informed a more accurate metagenomics binning of the most abundant microbes present.

Statistical analysis. Unless otherwise noted, statistics were done using the ecological statistics program PRIMER-e v7 (76). Metabolic data were normalized in PRIMER-e by dividing by sum total for each sample. The specific programs used in PRIMER-e were permutational multivariate analysis of variance (PERMANOVA) and distance-based linear models (DISTLM), the former of which calculates the significance and variance explained by a given factor and the latter determines which environmental variables correlate with the biological (microbiome) data. rfPermute (77), an R package for permuted random forests, was also performed to determine which annotated GC metabolites were indicative of microbial composition. PERMANOVA also partitions variance based on each factor, which is done in PRIMER-e by dividing the factor estimate by the sum total estimates of components of variation (ECoV). Traditional R^2 values were also calculated by dividing the sum of squares by total sum of squares. LMEs were carried out as described above; R^2 values for linear mixed models were calculated using the MuMIn package in R (78). Relate tests (analogous to Mantel tests) were used to compare GC-MS and LC-MS data. Bray-Curtis distances were used for all distance-based analyses. To consider repeated measures, linear mixed-effects modeling (nlme package in R) was used to analyze the stability of the microbiome and metabolome through time.

Ethics approval and consent to participate. This study utilized a subset of samples from a larger, longitudinal prospective cohort study designed to analyze the relationship between maternal stress and infant development conducted at the University of California, Irvine (UCI) (57). The University of California's Institutional Review Board approved the protocol, and written, informed consent was obtained from each participant. Research on human subjects was performed in accordance with the Declaration of Helsinki.

Data availability. Sequence data for 16S, ITS2, and shotgun metagenomes were deposited on the National Center for Biotechnology Information (NCBI) sequence read archive (SRA) under the BioProject accession number [PRJNA612083](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA612083). Metabolomics data for all samples can be found in Table S1. R scripts for statistical analysis are published on GitHub at <https://github.com/aoliver44/Cervicovaginal-Paper>.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, PDF file, 0.4 MB.

FIG S2, PDF file, 0.2 MB.

FIG S3, PDF file, 0.4 MB.

FIG S4, PDF file, 0.6 MB.

FIG S5, PDF file, 0.6 MB.

FIG S6, PDF file, 0.5 MB.

FIG S7, PDF file, 0.1 MB.

TABLE S1, XLSX file, 2.8 MB.

TABLE S2, XLSX file, 0.01 MB.

TABLE S3, XLSX file, 0.01 MB.

ACKNOWLEDGMENTS

The pilot grant awarded to start a UC Center for Pediatric Microbiome Research awarded by the UCI Institute for Clinical and Translational Science (UL1 TR001414) gave rise to this project. A.O. was supported by an NIH T32 training grant (1T32AI14134601A1) from UC Irvine's training program in microbiology and infectious diseases.

We thank Dan Cooper for enthusiasm and strategic advice. We thank the members of the Whiteson Lab, especially Whitney England for insight into metagenomics analysis, and Joann Phan and Tara Gallagher for thoughtful comments during the preparation of the manuscript. Heather Maughan made insightful and clarifying edits that we are grateful for.

We declare that we have no competing financial interests.

REFERENCES

- Smith MI, Yatsunenkov T, Manary MJ, Trehan I, Mkakosya R, Cheng J, Kau AL, Rich SS, Concannon P, Mychaleckyj JC, Liu J, Houpt E, Li JV, Holmes E, Nicholson J, Knights D, Ursell LK, Knight R, Gordon JI. 2013. Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. *Science* (80-) 339:548–554. <https://doi.org/10.1126/science.1229000>.
- Arrieta M-C, Stiemsma LT, Dimitriu PA, Thorson L, Russell S, Yurist-Doutsch S, Kuzeljevic B, Gold MJ, Britton HM, Lefebvre DL, Subbarao P, Mandhane P, Becker A, McNagny KM, Sears MR, Kollmann T, Mohn WW, Turvey SE, Finlay BB, CHILD Study Investigators. 2015. Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci Transl Med* 7:307ra152. <https://doi.org/10.1126/scitranslmed.aab2271>.
- Durack J, Kimes NE, Lin DL, Rauch M, McKean M, McCauley K, Panzer AR, Mar JS, Cabana MD, Lynch SV. 2018. Delayed gut microbiota development in high-risk for asthma infants is temporarily modifiable by Lactobacillus supplementation. *Nat Commun* 9:707. <https://doi.org/10.1038/s41467-018-03157-4>.
- Fujimora KE, Sitarik AR, Havstad S, Lin DL, Levan S, Fadrosch D, Panzer AR, LaMere B, Rackaityte E, Lukacs NW, Wegienka G, Boushey HA, Ownby DR, Zoratti EM, Levin AM, Johnson CC, Lynch SV. 2016. Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat Med* 22:1187–1191. <https://doi.org/10.1038/nm.4176>.
- Ege MJ, Bieli C, Frei R, van Strien RT, Riedler J, Ublagger E, Schram-Bijkerk D, Brunekreef B, van Hage M, Scheynius A, Pershagen G, Benz MR, Lauener R, von Mutius E, Braun-Fahrlander C, Parsifal Study Team. 2006. Prenatal farm exposure is related to the expression of receptors of the innate immunity and to atopic sensitization in school-age children. *J Allergy Clin Immunol* 117:817–823. <https://doi.org/10.1016/j.jaci.2005.12.1307>.
- Gomez de Agüero M, Ganal-Vonarburg SC, Fuhrer T, Rupp S, Uchimura Y, Li H. 2016. The maternal microbiota drives early postnatal innate immune development. *Science* 351:1296–1302. <https://doi.org/10.1126/science.aad2571>.
- Jašarević E, Howerton CL, Howard CD, Bale TL. 2015. Alterations in the vaginal microbiome by maternal stress are associated with metabolic reprogramming of the offspring gut and brain. *Endocrinology* 156:3265–3276. <https://doi.org/10.1210/en.2015-1177>.
- Walther-Antônio MRS, Jeraldo P, Berg Miller ME, Yeoman CJ, Nelson KE, Wilson BA, White BA, Chia N, Creedon DJ. 2014. Pregnancy's stronghold on the vaginal microbiome. *PLoS One* 9:e98514. <https://doi.org/10.1371/journal.pone.0098514>.
- DiGiulio DB, Callahan BJ, McMurdie PJ, Costello EK, Lyell DJ, Robaczewska A, Sun CL, Goltsman DSA, Wong RJ, Shaw G, Stevenson DK, Holmes SP, Relman DA. 2015. Temporal and spatial variation of the human microbiota during pregnancy. *Proc Natl Acad Sci U S A* 112:11060–11065. <https://doi.org/10.1073/pnas.1502875112>.
- Nasioudis D, Forney LJ, Schneider GM, Gliniewicz K, France M, Boester A, Sawai M, Scholl J, Witkin SS. 2017. Influence of pregnancy history on the vaginal microbiome of pregnant women in their first trimester. *Sci Rep* 7:10201. <https://doi.org/10.1038/s41598-017-09857-z>.
- Noyes N, Cho K-C, Ravel J, Forney LJ, Abdo Z. 2018. Associations between sexual habits, menstrual hygiene practices, demographics and the vaginal microbiome as revealed by Bayesian network analysis. *PLoS One* 13:e0191625. <https://doi.org/10.1371/journal.pone.0191625>.
- Gajer P, Brotman RM, Bai G, Sakamoto J, Schütte UME, Zhong X, Koenig SSK, Fu L, Ma ZS, Zhou X, Abdo Z, Forney LJ, Ravel J. 2012. Temporal dynamics of the human vaginal microbiota. *Sci Transl Med* 4:132ra52. <https://doi.org/10.1126/scitranslmed.3003605>.
- Serrano MG, Parikh HI, Brooks JP, Edwards DJ, Arodz TJ, Edupuganti L, Huang B, Girerd PH, Bokhari YA, Bradley SP, Brooks JL, Dickinson MR, Drake JI, Duckworth RA, Fong SS, Glascock AL, Jean S, Jimenez NR, Khoury J, Koparde VN, Lara AM, Lee V, Matveyev AV, Milton SH, Mistry SD, Rozycki SK, Sheth NU, Smirnova E, Vivadelli SC, Wijesooriya NR, Xu J, Xu P, Chaffin DO, Sexton AL, Gravett MG, Rubens CE, Hendricks-Muñoz KD, Jefferson KK, Strauss JF, Fettweis JM, Buck GA. 2019. Racioethnic diversity in the dynamics of the vaginal microbiome during pregnancy. *Nat Med* 25:1001–1011. <https://doi.org/10.1038/s41591-019-0465-8>.
- Petrova MI, Reid G, Vanechoutte M, Lebeer S. 2017. Lactobacillus iners: friend or foe? *Trends Microbiol* 25:182–191. <https://doi.org/10.1016/j.tim.2016.11.007>.
- Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, Karlebach S, Gortler R, Russell J, Tacket CO, Brotman RM, Davis CC, Ault K, Peralta L, Forney LJ. 2011. Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci* 108:4680–4687. <https://doi.org/10.1073/pnas.1002611107>.
- O'Hanlon DE, Moench TR, Cone RA. 2011. In vaginal fluid, bacteria associated with bacterial vaginosis can be suppressed with lactic acid but not hydrogen peroxide. *BMC Infect Dis* 11:200. <https://doi.org/10.1186/1471-2334-11-200>.
- Rampersaud R, Planet PJ, Randis TM, Kulkarni R, Aguilar JL, Lehrer RI, Ratner AJ. 2011. Inerolysin, a cholesterol-dependent cytolysin produced by Lactobacillus iners. *J Bacteriol* 193:1034–1041. <https://doi.org/10.1128/JB.00694-10>.
- Macklaim JM, Gloor GB, Anukam KC, Cribby S, Reid G. 2011. At the crossroads of vaginal health and disease, the genome sequence of Lactobacillus iners AB-1. *Proc Natl Acad Sci U S A* 108(Suppl):4688–4695. <https://doi.org/10.1073/pnas.1000086107>.
- Verstraelen H, Verhelst R, Claeys G, De Backer E, Temmerman M, Vanechoutte M. 2009. Longitudinal analysis of the vaginal microflora in pregnancy suggests that L. crispatus promotes the stability of the normal vaginal microflora and that L. gasseri and/or L. iners are more conducive to the occurrence of abnormal vaginal microflora. *BMC Microbiol* 9:116. <https://doi.org/10.1186/1471-2180-9-116>.
- Sobel JD. 2000. Bacterial vaginosis. *Annu Rev Med* 51:349–356. <https://doi.org/10.1146/annurev.med.51.1.349>.
- Fettweis JM, Serrano MG, Brooks JP, Edwards DJ, Girerd PH, Parikh HI, Huang B, Arodz TJ, Edupuganti L, Glascock AL, Xu J, Jimenez NR, Vivadelli SC, Fong SS, Sheth NU, Jean S, Lee V, Bokhari YA, Lara AM, Mistry SD, Duckworth RA, Bradley SP, Koparde VN, Orenda XV, Milton SH, Rozycki SK, Matveyev AV, Wright ML, Huzurbazar SV, Jackson EM, Smirnova E, Korlach J, Tsai Y-C, Dickinson MR, Brooks JL, Drake JI, Chaffin DO,

- Sexton AL, Gravett MG, Rubens CE, Wijesooriya NR, Hendricks-Muñoz KD, Jefferson KK, Strauss JF, Buck GA. 2019. The vaginal microbiome and preterm birth. *Nat Med* 25:1012–1021. <https://doi.org/10.1038/s41591-019-0450-2>.
22. Watts DH, Krohn MA, Hillier SL, Eschenbach DA. 1990. Bacterial vaginosis as a risk factor for post-caesarean endometritis. *Obstet Gynecol* 75:52–58.
 23. Acobsson BJ, Ernevi PP, Hidekel LC, Jörgen Platz-Christensen J. 2002. Bacterial vaginosis in early pregnancy may predispose for preterm birth and postpartum endometritis. *Acta Obstet Gynecol Scand* 81: 1006–1010. <https://doi.org/10.1034/j.1600-0412.2002.811103.x>.
 24. Hay PE, Lamont RF, Taylor-Robinson D, Morgan DJ, Ison C, Pearson J. 1994. Abnormal bacterial colonisation of the genital tract and subsequent preterm delivery and late miscarriage. *BMJ* 308:295–298. <https://doi.org/10.1136/bmj.308.6924.295>.
 25. Llahí-Camp JM, Rai R, Ison C, Regan L, Taylor-Robinson D. 1996. Association of bacterial vaginosis with a history of second trimester miscarriage. *Hum Reprod* 11:1575–1578. <https://doi.org/10.1093/oxfordjournals.humrep.a019440>.
 26. Donders GG, Van Bulck B, Caudron J, Londers L, Vereecken A, Spitz B. 2000. Relationship of bacterial vaginosis and mycoplasmas to the risk of spontaneous abortion. *Am J Obstet Gynecol* 183:431–437. <https://doi.org/10.1067/mob.2000.105738>.
 27. Ralph SG, Rutherford AJ, Wilson JD. 1999. Influence of bacterial vaginosis on conception and miscarriage in the first trimester: cohort study. *BMJ* 319:220–223. <https://doi.org/10.1136/bmj.319.7204.220>.
 28. Lev-Sagie A, Goldman-Wohl D, Cohen Y, Dori-Bachash M, Leshem A, Mor U, Strahilevitz J, Moses AE, Shapiro H, Yagel S, Elinav E. 2019. Vaginal microbiome transplantation in women with intractable bacterial vaginosis. *Nat Med* 25:1500–1504. <https://doi.org/10.1038/s41591-019-0600-6>.
 29. Masfari AN, Duerden BI, Kinghorn GR. 1986. Quantitative studies of vaginal bacteria. *Genitourin Med* 62:256–263.
 30. Pielou EC. 1966. The measurement of diversity in different types of biological collections. *J Theor Biol* 13:131–144. [https://doi.org/10.1016/0022-5193\(66\)90013-0](https://doi.org/10.1016/0022-5193(66)90013-0).
 31. MacIntyre DA, Chandramani M, Lee YS, Kindinger L, Smith A, Angelopoulos N. 2015. The vaginal microbiome during pregnancy and the postpartum period in a European population. *Sci Rep* 5:8988. <https://doi.org/10.1038/srep08988>.
 32. Brooks JP, Buck GA, Chen G, Diao L, Edwards DJ, Fettweis JM, Huzurbazar S, Rakitin A, Satten GA, Smirnova E, Waks Z, Wright ML, Yanover C, Zhou Y-H. 2017. Changes in vaginal community state types reflect major shifts in the microbiome. *Microb Ecol Health Dis* 28:1303265. <https://doi.org/10.1080/16512235.2017.1303265>.
 33. Grigoriou O, Baka S, Makrakis E, Hassiakos D, Kapparos G, Kouskouni E. 2006. Prevalence of clinical vaginal candidiasis in a university hospital and possible risk factors. *Eur J Obstet Gynecol Reprod Biol* 126:121–125. <https://doi.org/10.1016/j.ejogrb.2005.09.015>.
 34. Wang S, Wang Q, Yang E, Yan L, Li T, Zhuang H. 2017. Antimicrobial compounds produced by vaginal *Lactobacillus crispatus* are able to strongly inhibit *Candida albicans* growth, hyphal formation and regulate virulence-related gene expressions. *Front Microbiol* 8:564. <https://doi.org/10.3389/fmicb.2017.00564>.
 35. Narayanan TK, Rao GR. 1976. Beta-indoleethanol and beta-indolelactic acid production by *Candida* species: their antibacterial and autoantibiotic action. *Antimicrob Agents Chemother* 9:375–380. <https://doi.org/10.1128/aac.9.3.375>.
 36. Naz S, Cretenet M, Vernoux JP. 2013. Current knowledge on antimicrobial metabolites produced from aromatic amino acid metabolism in fermented products. In Méndez-Vilas A (ed), *Microbial pathogens and strategies for combating them: science, technology and education*, p 337–346. Formatex Research Centre, Badajoz.
 37. Esser C, Rannug A. 2015. The aryl hydrocarbon receptor in barrier organ physiology, immunology, and toxicology. *Pharmacol Rev* 67:259–279. <https://doi.org/10.1124/pr.114.009001>.
 38. Zelante T, Iannitti RG, Cunha C, De Luca A, Giovannini G, Pieraccini G, Cecchi R, D'Angelo C, Massi-Benedetti C, Fallarino F, Carvalho A, Puccetti P, Romani L. 2013. Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. *Immunity* 39:372–385. <https://doi.org/10.1016/j.immuni.2013.08.003>.
 39. Yang J, Chawla R, Rhee KY, Gupta R, Manson MD, Jayaraman A. 2020. Biphasic chemotaxis of *Escherichia coli* to the microbiota metabolite indole. *Proc Natl Acad Sci* 117:6114–6120. <https://doi.org/10.1073/pnas.1916974117>.
 40. Yeoman CJ, Yildirim S, Thomas SM, Durkin AS, Torralba M, Sutton G, Buhay CJ, Ding Y, Dugan-Rocha SP, Muzny DM, Qin X, Gibbs RA, Leigh SR, Stumpf R, White BA, Highlander SK, Nelson KE, Wilson BA. 2010. Comparative genomics of *Gardnerella vaginalis* strains reveals substantial differences in metabolic and virulence potential. *PLoS One* 5:e12411. <https://doi.org/10.1371/journal.pone.0012411>.
 41. Pybus V, Onderdonk AB. 1997. Evidence for a commensal, symbiotic relationship between *Gardnerella vaginalis* and *Prevotella bivia* involving ammonia: potential significance for bacterial vaginosis. *J Infect Dis* 175:406–413. <https://doi.org/10.1093/infdis/175.2.406>.
 42. Anahtar MN, Byrne EH, Doherty KE, Bowman BA, Yamamoto HS, Soumillon M, Padavattan N, Ismail N, Moodley A, Sabatini ME, Ghebremichael MS, Nusbaum C, Huttenhower C, Virgin HW, Ndung'u T, Dong KL, Walker BD, Fichorova RN, Kwon DS. 2015. Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. *Immunity* 42:965–976. <https://doi.org/10.1016/j.immuni.2015.04.019>.
 43. Forney LJ, Gliniewicz KS. 25 October 2018, publication date. Use of estrogenic compounds to manipulate the bacterial composition of vaginal communities. US patent number US2018305742.
 44. Huang L, Martin SM, Villanueva J, Greene S, Arehart K, Sayre C. 18 May 2006, publication date. Therapeutic agents for inhibiting and/or treating vaginal infection. US patent number US20060105963A1.
 45. Wu N, Zhang X, Li F, Zhang T, Gan Y, Li J. 2015. Spray-dried powders enhance vaginal siRNA delivery by potentially modulating the mucus molecular sieve structure. *Int J Nanomedicine (Lond)* 10:5383–5396. <https://doi.org/10.2147/IJN.S87978>.
 46. Eun Chang C, Kim S-C, So J-S, Shik Yun H. 2001. Cultivation of *Lactobacillus crispatus* KLB46 Isolated from human vagina. *Biotechnol Bioprocess Eng* 6:128–132. <https://doi.org/10.1007/BF02931958>.
 47. Van Der Linden PJQ, Wiegerinck MAHM. 1992. Cyclic changes in the concentration of glucose and fructose in human cervical mucus. *Fertil Steril* 57:573–577. [https://doi.org/10.1016/s0015-0282\(16\)54902-4](https://doi.org/10.1016/s0015-0282(16)54902-4).
 48. Donnarumma G, Molinaro A, Cimini D, De Castro C, Valli V, De Gregorio V, De Rosa M, Schiraldi C. 2014. *Lactobacillus crispatus* L1: high cell density cultivation and exopolysaccharide structure characterization to highlight potentially beneficial effects against vaginal pathogens. *BMC Microbiol* 14:137. <https://doi.org/10.1186/1471-2180-14-137>.
 49. Ojala T, Kankainen M, Castro J, Cerca N, Edelman S, Westerlund-Wikström B, Paulin L, Holm L, Auvinen P. 2014. Comparative genomics of *Lactobacillus crispatus* suggests novel mechanisms for the competitive exclusion of *Gardnerella vaginalis*. *BMC Genomics* 15:1070. <https://doi.org/10.1186/1471-2164-15-1070>.
 50. Wisselink H, Weusthuis R, Eggink G, Hugenholtz J, Grobden G. 2002. Mannitol production by lactic acid bacteria: a review. *Int Dairy J* 12: 151–161. [https://doi.org/10.1016/S0958-6946\(01\)00153-4](https://doi.org/10.1016/S0958-6946(01)00153-4).
 51. Wisselink HW, Moers APHA, Mars AE, Hoefnagel MHN, De Vos WM, Hugenholtz J. 2005. Overproduction of heterologous mannitol 1-phosphatase: a key factor for engineering mannitol production by *Lactococcus lactis*. *Appl Environ Microbiol* 71:1507–1514. <https://doi.org/10.1128/AEM.71.3.1507-1514.2005>.
 52. Gomez A, Nelson KE. 2018. The oral microbiome of children: development, disease, and implications beyond oral health. *Hum Microb* 73: 492–503. <https://doi.org/10.1007/s00248-016-0854-1>.
 53. Loening-Baucke V. 2005. Prevalence, symptoms and outcome of constipation in infants and toddlers. *J Pediatr* 146:359–363. <https://doi.org/10.1016/j.jpeds.2004.10.046>.
 54. Adachi S, Patton S. 1961. Presence and significance of lactulose in milk products: a review. *J Dairy Sci* 44:1375–1393. [https://doi.org/10.3168/jds.S0022-0302\(61\)89899-8](https://doi.org/10.3168/jds.S0022-0302(61)89899-8).
 55. Scalabre A, Jobard E, Demède D, Gaillard S, Pontoizeau C, Mouriquand P, Elena-Herrmann B, Mure P-Y. 2017. Evolution of newborns' urinary metabolomic profiles according to age and growth. *J Proteome Res* 16:3732–3740. <https://doi.org/10.1021/acs.jproteome.7b00421>.
 56. Angeletti S, Ciccozzi M. 2019. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry in clinical microbiology: an updating review. *Infect Genet Evol* 76:104063. <https://doi.org/10.1016/j.meegid.2019.104063>.
 57. Lindsay KL, Hellmuth C, Uhl O, Buss C, Wadhwa PD, Koletzko B, Entringer S. 2015. Longitudinal metabolomic profiling of amino acids and lipids across healthy pregnancy. *PLoS One* 10:e0145794. <https://doi.org/10.1371/journal.pone.0145794>.
 58. Entringer S, Buss C, Rasmussen JM, Lindsay K, Gillen DL, Cooper DM, Wadhwa PD. 2017. Maternal cortisol during pregnancy and infant

- adiposity: a prospective investigation. *J Clin Endocrinol Metab* 102: 1366–1374. <https://doi.org/10.1210/jc.2016-3025>.
59. Fiehn O. 2016. Metabolomics by gas chromatography-mass spectrometry: combined targeted and untargeted profiling. *Curr Protoc Mol Biol* 114: 30.4.1–30.4.32. <https://doi.org/10.1002/0471142727.mb3004s114>.
 60. Cajka T, Fiehn O. 2016. Toward merging untargeted and targeted methods in mass spectrometry-based metabolomics and lipidomics. *Anal Chem* 88:524–545. <https://doi.org/10.1021/acs.analchem.5b04491>.
 61. Cajka T, Smilowitz JT, Fiehn O. 2017. Validating quantitative untargeted lipidomics across nine liquid chromatography–high-resolution mass spectrometry platforms. *Anal Chem* 89:12360–12368. <https://doi.org/10.1021/acs.analchem.7b03404>.
 62. de Raad M, de Rond T, Rübél O, Keasling JD, Northen TR, Bowen BP. 2017. OpenMSI Arrayed Analysis Toolkit: analyzing spatially defined samples using mass spectrometry imaging. *Anal Chem* 89:5818–5823. <https://doi.org/10.1021/acs.analchem.6b05004>.
 63. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci* 108:4516–4522. <https://doi.org/10.1073/pnas.1000080107>.
 64. Baym M, Kryazhinskiy S, Lieberman TD, Chung H, Desai MM, Kishony RK. 2015. Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS One* 10:e0128036. <https://doi.org/10.1371/journal.pone.0128036>.
 65. Abarenkov K, Henrik Nilsson R, Larsson K-H, Alexander IJ, Eberhardt U, Erland S, Høiland K, Kjølner R, Larsson E, Pennanen T, Sen R, Taylor AFS, Tedersoo L, Ursing BM, Vrålstad T, Liimatainen K, Peintner U, Kõljalg U. 2010. The UNITE database for molecular identification of fungi - recent updates and future perspectives. *New Phytol* 186:281–285. <https://doi.org/10.1111/j.1469-8137.2009.03160.x>.
 66. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336. <https://doi.org/10.1038/nmeth.f.303>.
 67. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. 2019. nlme: linear and nonlinear mixed effects models. R package version 3.1-141. <https://CRAN.R-project.org/package=nlme>.
 68. R Core Team. 2018. R: a language and environment for statistical computing. Foundation for Statistical Computing, Vienna, Austria. <http://www.r-project.org/>.
 69. Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864. <https://doi.org/10.1093/bioinformatics/btr026>.
 70. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
 71. Franzosa EA, McIver LJ, Rahnward G, Thompson LR, Schirmer M, Weingart G, Lipson KS, Knight R, Caporaso JG, Segata N, Huttenhower C. 2018. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods* 15:962–968. <https://doi.org/10.1038/s41592-018-0176-y>.
 72. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. 2011. Metagenomic biomarker discovery and explanation. *Genome Biol* 12:R60. <https://doi.org/10.1186/gb-2011-12-6-r60>.
 73. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
 74. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3:e1319. <https://doi.org/10.7717/peerj.1319>.
 75. Menzel P, Ng KL, Krogh A. 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* 7:11257. <https://doi.org/10.1038/ncomms11257>.
 76. Clark KR, Warwick RM, Gorley RN, Somerfield PJ. 2014. Change in marine communities: an approach to statistical analysis and interpretation, 3rd edition. PRIMER-E, Plymouth, UK.
 77. Archer E. 2019. rfPermute: estimate permutation p-values for random forest importance metrics. <https://cran.r-project.org/package=rfPermute>.
 78. Barton K. 2019. MuMIn: multi-model inference.