Disentangling race, environment and the microbiome in a study of preterm birth risk

Shan Sun (ssun5@uncc.edu)
University of North Carolina at Charlotte https://orcid.org/0000-0003-0349-2664

Myrna Serrano
Virginia Commonwealth University

Jennifer Fettweis
Virginia Commonwealth University

Patricia Basta
University of North Carolina at Chapel Hill

Emma Rosen
University of North Carolina at Chapel Hill

Kim Ludwig
University of North Carolina at Chapel Hill

Alicia Sorgen
University of North Carolina at Charlotte

Ivory Blakley
University of North Carolina at Charlotte

Michael Wu
Fred Hutchinson Cancer Research Center https://orcid.org/0000-0002-3357-6570

Nancy Dole
University of North Carolina at Chapel Hill

John Thorp
University of North Carolina at Chapel Hill

Anna Siega-Riz
University of Massachusetts Amherst

Gregory Buck
Virginia Commonwealth University https://orcid.org/0000-0003-4621-3987

Anthony Fodor
UNC Charlotte

Stephanie Engel
University of North Carolina at Chapel Hill

Article

Keywords: vaginal microbiome, preterm birth, race, douching

Posted Date: March 2nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-228078/v1

Abstract

Previous studies have investigated the associations between the vaginal microbiome and preterm birth (PTB), with the aim of determining whether differences in community patterns meaningfully alter risk, and could therefore be the target of intervention. We report on vaginal microbial analysis on a subset of the Pregnancy, Infection, and Nutrition (PIN) Study, a prospectively enrolled cohort of women in central North Carolina between 1995-2001. We selected a nested case-control subset of this cohort, including 464 White women (375 term birth and 89 spontaneous PTB, sPTB) and 360 Black women (276 term birth and 84 sPTB). Microbial DNA was extracted from genital track swabs collected mid-pregnancy, and subjected to 16S rRNA taxonomic profiling. We found that microbial community structure is associated with race and sPTB, although the influence of race is stronger than the influence of sPTB. The microbiome of Black women has higher alpha-diversity, higher abundance of Lactobacillus iners and lower abundance of Lactobacillus crispatus. These differences were obscured once maternal douching behavior was considered—specifically, among women who douche, there were no significant differences in microbiome by race. The sPTB associated microbiome exhibited a lower abundance of L. crispatus, while alpha diversity and L. iners were not significantly different. Associations between the microbiome and sPTB were only significant in women who do not douche. While race was a strong predictor of microbial community structure, we also observed strong intercorrelations between a range of maternal factors, including poverty, education, marital status, age, douching and race, with microbiome effect sizes in the range of 1.8-5.2% in univariate models. Therefore, race may simply be a proxy for other socially driven factors that differentiate microbiome community structures. Future work will continue to refine reliable microbial biomarkers for preterm birth across diverse cohorts.

1.0 Introduction

Over 10% of all pregnancies in the US are preterm, almost 16% among Black women [1]. Intrauterine infection is widely speculated to underlie some portion of preterm deliveries; however, the benefit of prophylactic antibiotic therapy for the prevention of preterm birth (PTB) is not universal, and appears to depend at least in part on timing of treatment in pregnancy, route (oral versus IV) of antibiotic administration, and clinical presentation (*e.g.*, preterm premature rupture of membranes versus intact membranes) [2–5], which may reflect etiologic heterogeneity among preterm births. One pathway through which pathogenic microorganisms may gain access to the amniotic cavity is by ascending from the vagina and the cervix [6].

The vaginal microbiome represents a physiological barrier to this route, generally through “healthy” bacteria (mostly but not always of the *Lactobacillus* genus) producing lactic acid and lowering the vaginal pH [7]. However, there is a tremendous diversity in the species of *Lactobacillus* present in the vagina, and these species may produce varying levels of lactic acid and have different tolerance for anaerobic members of the microbial community [8]. Moreover, some relatively common vaginal microbiome profiles contain few *Lactobacillus* spp. [9,10]. Vaginal microbiome lacking *Lactobacillus*

species (spp.) of any type tend to have higher pH, more taxon diversity, and often prominently represent organisms that make up the bacterial vaginosis (BV) diagnosis by Nugent Score [7,10,11]. This profile, which is associated with adverse outcomes, tends to occur more frequently in Black women [10–12].

In a recent large study, Fettweis *et al.* described vaginal microbial community differences among 1,268 African American and 416 European American women [11]. They report significant differences in the microbiome profiles between these groups, with African American women having greater microbial diversity and dominance by BV-associated organisms. In a longitudinal study of pregnant women, Fettweis *et al.* found that the abundance of *L. crispatus* was reduced in PTB groups and the abundances of *Candidatus* Lachnocurva vaginae (BVAB1), *Prevotella* cluster 2, *Sneathia amnii* and several additional taxa were increased [13]. Interestingly, these signals were strongest earlier in gestation and were largely driven by samples from women of African ancestry. Callahan *et al.* analyzed the associations between vaginal microbiome and PTB in two racially distinct cohorts, with one being a low risk predominantly Caucasian cohort (n = 39) and the other one a high risk majorly African American cohort (n = 96) [14]. They found that *L. crispatus* but not *L. iners* was associated with low PTB risk in both cohorts. The lower abundance of *L. jensenii* and *L. gasseri* was associated with PTB only in the high-risk cohort, while the higher abundance of *G. vaginalis* was associated with PTB only in the low-risk cohort. Kindinger *et al.* studied the vaginal microbiome in a cohort of women at risk of preterm birth (total n = 161: Black n = 30, Caucasian n = 104, and Asian n = 27) and found that *L. iners* dominance is associated with higher PTB risk and *L. crispatus* is associated with lower PTB risk [15]. However, no significant difference was identified when stratified by race, perhaps due to the small sample size.

The present literature appears to support that Black women often have more diverse vaginal microbial community patterns [10,11] and that women with higher abundance of *L. crispatus* have overall lower risk of PTB [13–15]. What is not yet clear, given the known racial disparity of PTB [16,17], is whether microbial community patterns are associated with PTB independent of racial differences in microbial community structure. To answer this question, we utilized the Pregnancy, Infection, and Nutrition (PIN) Study, a prospectively enrolled pregnancy cohort of women in central North Carolina, to investigate the relationship between second trimester vaginal microbial community patterns and PTB, and to disentangle differences in association by race. This cohort of low-risk women represents the ideal setting to answer this question given the rich social, behavioral, demographic and clinical data assembled that characterize in detail known determinants of PTB [18].

2.0 Methods

2.1 Study Population

The Pregnancy, Infection and Nutrition (PIN) study enrolled pregnant women with singleton pregnancies in central North Carolina from August 1995 – Feb 2001. Women were recruited from prenatal clinics at the University of North Carolina Hospitals, Wake County Human Services and the Wake Area Health

Education Center. Eligibility criteria included gestational age at enrollment between 24–29 weeks', ability to communicate in English, age 16 years or older, access to telephone, and plans to deliver at the recruitment site [19]. This study was approved by the Institutional Review Board at UNC Chapel Hill.

At enrollment, women provided blood, urine and genital tract specimens. They were also randomly assigned to a "subcohort" for future nested case-cohort studies intending to employ detailed biological measurements that would be infeasible to conduct in the entire population. Because women were randomly assigned to this subset irrespective of pregnancy outcome, this subcohort should reflect the exposure distribution in the cohort as a whole. In the subsequent two weeks following enrollment, women completed a telephone interview that collected information on social, demographic, medical and behavioral risk factors for adverse pregnancy outcomes. In total, 3,163 women were recruited into the PIN study during this period. The current study was restricted to spontaneous preterm and subcohort members who self-classified as black or white race (n = 824), including 375 White women with term birth, 89 with spontaneous preterm birth, 276 Black women with term birth and 84 with spontaneous preterm birth. Complete selection criteria for the current analysis are described in Fig. S1.

2.2 Preterm Birth Clinical Presentation

Gestational age at delivery was assigned by early ultrasound (completed prior to 22-week gestation) in 90% of the population, or last menstrual period date if ultrasound was unavailable [20]. Preterm birth was defined as < 37 completed weeks' gestation. Preterm clinical presentation was determined by obstetrician review, and classified as preterm labor (PTL), preterm premature rupture of amniotic membranes (PPROM) in which membranes ruptured four or more hours before the onset of labor, and medically indicated. For the current study, we combined PTL and PPROM into a single clinical presentation of spontaneous preterm birth (sPTB). Among sPTB cases, gestational ages varied from 26–36 completed weeks.

2.3 Covariates

We selected covariates for consideration as predictors of the microbiome or confounders of the microbiome-PTB association based on prior literature. These included maternal age at enrollment, maternal education, marital status, pre-pregnancy weight and height to calculate body mass index (BMI), parity, any smoking during pregnancy, maternal household percent of poverty based on the 1996 census, douching before pregnancy, maternal self-reported depressive symptoms, and the number of negative life events. Maternal depressive symptoms were measured based on the Center for Epidemiological Studies Depression scale (CES-D) [21]. Negative life events were assessed by a modified Life-Events Inventory (LEI) [22].

2.4 DNA extraction and Sequencing

Swabs were collected between 24–29 weeks' gestation from the posterior vaginal apex, and stored at -70C. For the current study, swabs were thawed on ice and processed essentially as previously described [13]. In brief, DNA was extracted using the PowerSoil DNA Isolation Kit (Qiagen), eluted in 100 µL water and quantified using PicoGreen. Extracted DNA was amplified with barcoded primers targeting the V1–V3

hypervariable regions of the bacterial 16S rRNA gene using protocols established in the Vaginal Human Microbiome Project at VCU [11]. Samples were multiplexed (384 samples/run) using a sample-specific dual-index strategy and sequenced on Illumina MiSeq sequencers (2 x 300 base paired end protocol). The paired-end quality-aware raw sequence files were demultiplexed into sample-specific data, and merged and quality-filtering using MeFiT [23]. Samples with fewer than 1,000 high-quality reads were excluded. The sequences generated can be accessed at NCBI with BioProject ID PRJNA694098.

2.5 Bioinformatics and Statistical Analysis Approach

Comprehensive 16S rRNA gene-based taxonomic survey of the vaginal microbial profiles yielded a mean count of 43,276 reads/sample with minimum and maximum read counts of 1,824 and 186,784, respectively. Over 99.9% of the high-quality single reads generated overlapping pair-end reads. High-quality sequences were assigned to the species-level taxonomic assignments for vaginal samples using STIRRUPS [24], an analysis platform that employs the USEARCH algorithm [25] combined with a curated 16S rRNA sequence database. Paired reads which did not align to the same reference sequence were discarded as chimeras. Analyses with DADA2 and SILVA 132 release were used as alternative pipelines.

The PCoA ordinations were calculated based on the Bray-Curtis dissimilarity between samples with function ‘capscale’ and visualized with ‘ordiplot’ in R package ‘vegan’. PERMANOVA tests were used to analyze the associations between microbiome and host factors with function ‘adonis’ in the same package. Shannon index was used to calculate the alpha-diversity of microbial communities.

Two methods were applied to determine the vaginal microbial community states or vagitypes based on the taxonomic composition. First, vagitypes were determined by the dominant species with a relative abundance > 30%. The microbiome was characterized as ‘no type’ when the relative abundance of all species was lower than 30%. We also used the hierarchical cluster analysis with R function ‘hclust’ to confirm the existence of vagitypes. The associations between preterm/term birth, vagitypes and host factors were determined with Fisher’s exact test. The associations between species and preterm/term birth, race and douching were primarily analyzed with Wilcoxon tests. P-values were adjusted for multiple testing using the Benjamini-Hochberg method [26].

3.0 Results

3.1 Maternal characteristics, vaginal microbiome and spontaneous preterm birth

The vaginal microbiome profiles differed significantly by maternal race and spontaneous preterm birth (Fig. 1a and b). PCoA ordination of the microbiome showed a separation of the 95% confidence limits of Black and White women (Fig. 1a), with a PERMANOVA R^2 of 1.8%. However, PCoA ordination of the microbiome showed only a modest separation of the 95% confidence limits of spontaneous preterm and term birth groups (Fig. 1b), with a relatively small PERMANOVA R^2 (0.45%). A number of maternal features were significantly associated with vaginal microbiome, including percent of poverty level, years of education, marital status, age at mid-pregnancy, douching, self-reported depression, negative life

events, and parity (Fig. 1c), which are intercorrelated and differentially distributed by maternal race (Fig. S2); therefore, identifying the underlying causal attribute is challenging. Because of the existing literature documenting differences in community patterns across racial and ethnic populations [27], we orient our results according to maternal self-reported race; however, these patterns likely reflect a complex interplay between social and environmental factors for which race is a marker but not a causal factor. The microbiome of Black women has higher alpha-diversity, higher abundance of *L. iners* and lower abundance of *L. crispatus* (Fig. 1d). The spontaneous preterm birth associated microbiome has lower abundance of *L. crispatus*, while alpha diversity and *L. iners* were not significantly different (Fig. 1e).

3.2 Vagitypes, spontaneous preterm birth and race

The taxonomic composition suggested that most of the vaginal microbiomes were dominated by a single taxon, with the most prevalent species being *L. iners*, *L. crispatus*, *L. gasseri*, *L. jensenii*, Lachnospiraceae member BVAB1 (BVAB1 has been named provisionally: "*Candidatus* Lachnocurva vaginae") and *Gardnerella* spp. (Fig. 2a). Because of the discrete community structures of vaginal microbiome, we classified the microbiome into vagitypes based on the dominant species with relative abundance > 30% following previously reported methods [13]. The PCoA ordination of the microbiome showed that different vagitypes generally formed distinguishable clusters especially for those dominated by *L. iners* and *L. crispatus* (Fig. 2b). The same PCoA ordination but colored by race and spontaneous preterm birth showed that the microbiomes of White women, especially those who will experience term birth, are more likely to be of *L. crispatus* cluster (Fig. 2c).

To determine whether the vagitypes were associated with spontaneous preterm birth, we calculated and compared the percentage of spontaneous preterm birth cases in each vagitype (Fig. 2d). In this analysis, the vagitypes dominated by non-*Lactobacillus* were grouped as Others. The *Lactobacillus* vagitypes except *L. iners* and *L. crispatus* were grouped as Lacto_other in order to simplify the model and for sample size considerations. We found that the percentage of spontaneous preterm birth cases with *L. crispatus* vagitype was significantly lower than the other three types, with a spontaneous preterm birth percentage of 13% compared to 22%, 25% and 26% for *L. iners*, Lacto_other and Others respectively (Fisher's exact test) (Fig. 2d). Because this study oversampled the underlying cohort for preterm cases, the percentage of preterm birth above does not reflect the underlying risk in the population. When the oversampling of preterm birth was accounted for, the risk of spontaneous preterm birth across vagitypes followed the same pattern: 3.5%, 5.9%, 8.8% and 9.1% for vagitypes *L. crispatus*, *L. iners*, Lacto_other and Others respectively. The Shannon diversity of *L. crispatus* cluster was significantly lower than the other three vagitypes (Fig. 2e).

We analyzed whether the percentages of vagitypes were different between Black and White women, and found Black women had a higher percentage of the *L. iners* vagitype and lower percentages of *L. crispatus* and Lacto_other vagitypes as compared to White women. We next examined whether maternal race modified the association between the *L. crispatus* vagitype and preterm birth and found no difference between Black and White women (Fisher's exact test, OR = 1.11, CI = 0.41–2.88), and there is no

difference for *L. iners* between Black and White women as well (Fisher's exact test, OR = 0.94, CI = 0.56–1.57).

3.3 Microbiome, douching and spontaneous preterm birth

To better understand the relationship between maternal race, douching, and vaginal microbiome, we created 4 distinct groups for a subset of participants with douching information (n = 489): Black, No Douching (B_N, n = 78); Black, Douching (B_D, n = 110); White, No Douching (W_N, n = 199); and White, Douching (W_D, n = 102). PCoA ordination showed that the White non-douching group formed a separate cluster from the other three groups (Fig. 3a). This was supported by the PERMANOVA tests that indicated that the White non-douching group is significantly different from the Black non-douching group (R^2 = 0.0284, P = 0.001), while the two douching groups were not significantly different from each other (P = 0.401). This is consistent with the vagitype composition of these 4 groups, with non-douching White women associated with higher percentage of *L. crispatus*, Lacto_other and lower percentage of *L. iners* vagitypes (Fig. 3b). Additionally, the alpha diversity and the abundance of *L. iners* and *L. crispatus* (Fig. 3c) showed similarly that the White non-douching group was significantly different from the others (Wilcoxon test, P < 0.05).

At the individual species level, there were 23 taxa that showed significant associations with race in the non-douching participants but there was none significant by race in douching group (Fig. 3d). Likewise, there are 22 taxa associated with douching in White participants, but no taxa associated with douching in Black participants (Fig. 3e). Taken together, these data show that the association of douching and the microbiome was much stronger for White participants, while race was significantly associated with the variation of the vaginal microbiome only for non-douching participants.

We next examined the relationship of spontaneous preterm birth and the microbiome in the 4 groups. With PERMANOVA tests, the vaginal microbiome was significantly associated with spontaneous preterm birth only in the non-douching participants and for both race groups, while this association was not significant for the douching participants (Fig. 3f). Moreover, compared to the PERMANOVA test for pregnancy outcomes without stratification of race and douching, the effect sizes (PERMANOVA R^2) here are much increased, from 0.45–3% in Black women and 1% in White women.

4.0 Discussion

In this large, prospective pregnancy cohort, we analyzed the association between mid-pregnancy vaginal microbiome, race, and sPTB. We found that vaginal microbiomes were significantly associated with sPTB, race, douching and other maternal factors. Many of these maternal factors, like poverty, education, marital status, age, douching and race, have stronger associations with the vaginal microbiome than the vaginal microbiome has with sPTB (Fig. 1c). Consistent with previous studies [10,27], we found that the vaginal microbiomes of Black and White women were significantly different, with higher alpha diversity, higher abundance of *L. iners* and lower abundance of *L. crispatus* for Black women. The microbial difference between sPTB and term controls is mainly driven by a higher *L. crispatus* abundance in term

controls, similar to previous reports [13,14]. Because of the strong intercorrelations between maternal factors such as race, poverty, education, marital status and douching (Fig. S2), we stratified the dataset by race and douching with the aim of uncovering potentially stronger sPTB microbial signatures that are independent of race and douching.

With the community state types assigned based on the most abundant taxon, the sPTB risk associated with *L. crispatus* dominated community state is about 60% of that for *L. iners* dominated microbiome (Fig. 2d, 13% and 22% in all participants, and 3.5% and 5.9% when oversampling of cases in this nested case-control design is accounted for). The alpha diversity of *L. crispatus* dominated microbiome is significantly lower than that of *L. iners* dominated microbiome, indicating that *L. crispatus* may suppress the colonization and development of BV-like microbiome while *L. iners* does not. Compared to *L. crispatus*, vaginal microbiome dominated by *L. iners* also more often shift towards a diverse community [28]. For example, *L. iners* enhance the adhesion of *Gardnerella* spp. to cervical epithelial cells, and *Gardnerella* spp. displaced adherent *L. crispatus* but not *L. iners* from epithelial cells [29]. Previous research also suggests that *L. crispatus* and *Gardnerella* were exclusive while *L. iners* and *Gardnerella* often co-exist [14]. This may explain the higher sPTB risk associated with *L. iners* dominated microbiome compared to *L. crispatus* in our population. The sPTB risk associated with community state “Lacto_other” dominated by other *Lactobacillus* species (mostly *L. gasseri* and *L. jensenii/fornicalis/psittaci*) is also significantly higher than *L. crispatus* dominated community state, indicating that these species were also not as protective as *L. crispatus.*

With a similar number of Black and White participants, we analyzed the associations between microbiome and the sPTB risk in each race separately and found that risk of sPTB associated with *L. crispatus* and *L. iners* are similar for Black and White women (Fig. 2g). Although at the US population level, black women have substantially higher risk of PTB [17], in the PIN study specifically, black race is only marginally associated with PTB (OR 1.3, 95% CI 1.0, 1.6) [18]. Our findings, that race does not modify the association between *L. crispatus* and *L. iners* and sPTB, may suggest that the disparity in PTB rates at the population level may in part be due to the lower prevalence of *L. crispatus* dominated microbiome among Black women. While the associations between *L. crispatus, L. iners*, and sPTB are independent of race, we could not verify whether risk associated with other community patterns are also consistent between races, because of their relatively low prevalence among participants. Future studies with a larger number of participants are needed to further investigate other taxa.

Douching is often associated with BV [30–32], although it is difficult to determine whether douching increases the risk of BV or BV leads to douching. In this study, we found that douching played an important role in the structure of the vaginal microbiome. Among women who did not douche, Black and White women have different microbiome. Specifically, White women had a notably higher abundance of *L. crispatus*, lower abundance of *L. iners* and higher abundance of other *Lactobacillus* species. However, among women who did douche, the microbiomes of Black and White women were similar, and featured lower abundance of *L. crispatus* and higher abundance of *L. iners*. It was reported that the genome of *L.*

iners AB-1 contains genes that could contribute to its survival in an environment of fluctuating conditions, including Fe-S cluster protein for oxidative stress, alkaline shock and universal stress proteins [28]. *L. iners* also has a stronger ability to adhere to human fibronection than other vaginal bacteria strains such as *L. crispatus* ATCC 3800 [33]. Thus, it is possible that douching habits influence *L. iners* dominated microbiome less than *L. crispatus* dominated microbiome. At the same time, the percentage of sPTB cases was higher in douching groups in both Black and White women although causality cannot be directly inferred. It is possible that a pre-existing dysbiotic state caused both douching behavior and sPTB, or alternatively, that douching disrupted the healthier *L. crispatus* dominated microbiome that then shifted to a higher risk microbiome, ultimately leading to sPTB. Future studies with longitudinal vaginal track sampling, and longitudinal information on douching behavior, would be required to disentangle these interconnected features. Regardless, our results suggest that douching behavior is significantly associated with the vaginal microbiome, and that race-related differences in vaginal microbiome are erased in the population of women that report douching.

In summary, in this prospective study of mid-pregnancy microbiome and sPTB in a well characterized cohort of Black and White women, we found that the vaginal microbiome of Black women was characterized by higher diversity, lower abundance of *L. crispatus*, and higher abundance of *L. iners*. These differences were obscured once maternal douching behavior was considered—specifically, among women who douche, there were no material differences in microbiome by race. Additionally, we found that women with microbiome dominated by *L. crispatus* had lower risk of sPTB, and women with microbiome dominated by *L. iners* had higher risk of sPTB, and these associations were the same for Black and White women. To our knowledge, this is the first study of the vaginal microbiome and sPTB to consider the influence of douching, and we found that douching has a significant influence on the vaginal microbiome that should be considered in future studies.

Finally, it is important to note that while we present differences in microbial community patterns by race to be consistent with the prior literature [10,11], we observed strong inter-correlations across a number of maternal factors whose effects cannot easily be separated. These intercorrelated factors include race, poverty level, psychosocial stress, education, marital status, and maternal age. In this as with many other medical research studies, maternally self-classified race only crudely captures complex social determinants of health [34], and thus disparities in microbial community patterns that we observe in relation to race may actually result from factors such as but not limited to her diet, her access to high quality medical care, her social support and life experiences, her psychosocial stress, and her experiences of discrimination. These pathways, that may explain differences in vaginal microbial community patterns by race, need further investigation and elucidation. Although this is one of the largest studies of the associations between vaginal microbiome and sPTB, it still lacks power for analyzing less abundant microbes and whether the combination of race or other social factors and douching influence the consistency of microbial signatures. Pooled studies across cohorts with similar metagenomics data may enable a more precise investigation of rare species as well as the influence of maternal factors that may explain or modify effects of vaginal microbiome on sPTB.

Declarations

Funding

This study was funded in part by a grant from NIH/NIMHD (R01MD011504), NIH/NIEHS (P30 ES010126). ER was supported by T32ES007018. The Pregnancy, Infection, and Nutrition study was supported by Grants HD37584, HD39373, and DK61981. The General Clinic Research Center was supported by the National Institutes of Health General Clinical Research Centers program of the Division of Research Resources Grant RR00046.

Ethics approval and consent to participate

This study was approved by the Institutional Review Board at UNC Chapel Hill.

Availability of data and material

The datasets analyzed in this study are available at NCBI with BioProject PRJNA694098. R scripts used in this study are available at Github (https://github.com/ssun6/PINmicrobiome). Additional requests and questions can be addressed to SS.

Authors' contributions

SME, AAF and GAB contributed to all aspects of the study, including conception, design, data acquisition, analysis, and supervision. MGS, JMF, PB, KL, ND, JT, AMSR contributed to acquisition of data, interpretation of results. SS, JMF, ER, AAS, ICB, MCW, AAF and SME contributed to analysis and interpretation of data. All authors contributed to writing, review, and/or revision of the manuscript, and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

References

1. Martin, J. A., Hamilton, B. E. & Osterman, M. J. K. in *National vital statistics reports* Vol. 64 (National Center for Health Statistics, Hyattsville, MD, 2015).
2. Flenady, V., Hawley, G., Stock, O. M., Kenyon, S. & Badawi, N. Prophylactic antibiotics for inhibiting preterm labour with intact membranes. *Cochrane Database Syst Rev* **12**, CD000246, doi:10.1002/14651858.CD000246.pub2 (2013).
3. Kenyon, S., Boulvain, M. & Neilson, J. P. Antibiotics for preterm rupture of membranes. *Cochrane Database Syst Rev* **12**, CD001058, doi:10.1002/14651858.CD001058.pub3 (2013).
4. Oliver, R. S. & Lamont, R. F. Infection and antibiotics in the aetiology, prediction and prevention of preterm birth. *J Obstet Gynaecol* **33**, 768–775, doi:10.3109/01443615.2013.842963 (2013).

5. Mercer, B. Antibiotics in the management of PROM and preterm labor. *Obstet Gynecol Clin North Am* **39**, 65–76, doi:10.1016/j.ogc.2011.12.007S0889-8545(11)00130-6 [pii] (2012).
6. Goldenberg, R. L., Culhane, J. F., Iams, J. D. & Romero, R. Epidemiology and causes of preterm birth. *Lancet* **371**, 75–84, doi:10.1016/S0140-6736(08)60074-4 S0140-6736(08)60074-4 [pii] (2008).
7. Brotman, R. M. Vaginal microbiome and sexually transmitted infections: an epidemiologic perspective. *J Clin Invest* **121**, 4610–4617, doi:10.1172/JCI5717257172 [pii] (2011).
8. Ma, B., Forney, L. J. & Ravel, J. Vaginal microbiome: rethinking health and disease. *Annu Rev Microbiol* **66**, 371–389, doi:10.1146/annurev-micro-092611-150157 (2012).
9. Hickey, R. J., Zhou, X., Pierson, J. D., Ravel, J. & Forney, L. J. Understanding vaginal microbiome complexity from an ecological perspective. *Transl Res* **160**, 267–282, doi:10.1016/j.trsl.2012.02.008 S1931-5244(12)00070 – 9 [pii] (2012).
10. Ravel, J. *et al.* Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A* **108 Suppl 1**, 4680–4687, doi:10.1073/pnas.1002611107 1002611107 [pii] (2011).
11. Fettweis, J. M. *et al.* Differences in vaginal microbiome in African American women versus women of European ancestry. *Microbiology* **160**, 2272 (2014).
12. Hyman, R. W. *et al.* Diversity of the vaginal microbiome correlates with preterm birth. *Reprod Sci* **21**, 32–40, doi:10.1177/1933719113488838 1933719113488838 [pii] (2014).
13. Fettweis, J. M. *et al.* The vaginal microbiome and preterm birth. *Nature medicine* **25**, 1012–1021 (2019).
14. Callahan, B. J. *et al.* Replication and refinement of a vaginal microbial signature of preterm birth in two racially distinct cohorts of US women. *Proceedings of the National Academy of Sciences* **114**, 9966–9971 (2017).
15. Kindinger, L. M. *et al.* The interaction between vaginal microbiota, cervical length, and vaginal progesterone treatment for preterm birth risk. *Microbiome* **5**, 6 (2017).
16. MacDorman, M. F. in *Seminars in perinatology.* 200–208 (Elsevier).
17. Reagan, P. B. & Salsberry, P. J. Race and ethnic differences in determinants of preterm birth in the USA: broadening the social context. *Social science & medicine* **60**, 2217–2228 (2005).
18. Savitz, D. A. *et al.* Behavioral influences on preterm birth: integrated analysis of the pregnancy, infection, and nutrition study. *Maternal and child health journal* **16**, 1151–1163 (2012).
19. Savitz, D. *et al.* Determinants of participation in an epidemiological study of preterm delivery. *Paediatric and Perinatal Epidemiology* **13**, 114–125 (1999).
20. Savitz, D. A. *et al.* Comparison of pregnancy dating by last menstrual period, ultrasound scanning, and their combination. *American journal of obstetrics and gynecology* **187**, 1660–1666 (2002).
21. Radloff, L. S. The CES-D scale: A self-report depression scale for research in the general population. *Applied psychological measurement* **1**, 385–401 (1977).
22. Sarason, I. G., Johnson, J. H. & Siegel, J. M. Assessing the impact of life changes: development of the Life Experiences Survey. *Journal of consulting and clinical psychology* **46**, 932 (1978).

23. Parikh, H. I., Koparde, V. N., Bradley, S. P., Buck, G. A. & Sheth, N. U. MeFiT: merging and filtering tool for illumina paired-end reads for 16S rRNA amplicon sequencing. *BMC bioinformatics* **17**, 1–6 (2016).
24. Fettweis, J. M. *et al.* Species-level classification of the vaginal microbiome. *BMC genomics* **13**, 1–9 (2012).
25. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
26. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* **57**, 289–300 (1995).
27. Serrano, M. G. *et al.* Racioethnic diversity in the dynamics of the vaginal microbiome during pregnancy. *Nature medicine* **25**, 1001–1011 (2019).
28. Petrova, M. I., Reid, G., Vaneechoutte, M. & Lebeer, S. Lactobacillus iners: friend or foe? *Trends in microbiology* **25**, 182–191 (2017).
29. Castro, J. *et al.* Reciprocal interference between Lactobacillus spp. and Gardnerella vaginalis on initial adherence to epithelial cells. *International journal of medical sciences* **10**, 1193 (2013).
30. Brotman, R. M. *et al.* A longitudinal study of vaginal douching and bacterial vaginosis—a marginal structural modeling analysis. *American journal of epidemiology* **168**, 188–196 (2008).
31. Holzman, C. *et al.* Factors linked to bacterial vaginosis in nonpregnant women. *American journal of public health* **91**, 1664–1670 (2001).
32. Hutchinson, K. B., Kip, K. E. & Ness, R. B. Vaginal douching and development of bacterial vaginosis among women with normal and abnormal vaginal microflora. *Sexually transmitted diseases* **34**, 671–675 (2007).
33. McMillan, A., Macklaim, J. M., Burton, J. P. & Reid, G. Adhesion of Lactobacillus iners AB-1 to human fibronectin: a key mediator for persistence in the vagina? *Reproductive Sciences* **20**, 791–796 (2013).
34. Fontanarosa, P. B. & Bauchner, H. Race, ancestry, and medical research. *Jama* **320**, 1539–1540 (2018).

Tables

Table 1. Characteristics of Study Population, The Pregnancy Infection and Nutrition Cohort, 1995-2001 (n = 824)

Covariate	White		Black	
	Term (N=375) N, % or Mean (SD)	Spontaneous Preterm (N=89) N, % or Mean (SD)	Term (N=276) N, % or Mean (SD)	Spontaneous Preterm (N=84) N, % or Mean (SD)
Maternal age (years)	27.4 (5.83)	26.9 (7.17)	24.1 (5.36)	25.4 (5.83)
Maternal education (years)	14.1 (3.32)	13.1 (2.82)	12.4 (2.00) Missing = 1	12.7 (1.67)
Pre-pregnancy BMI (kg/m^2)	25.6 (6.93) Missing = 7	25.2 (6.56) Missing = 2	27.4 (7.99) Missing = 9	27.7 (7.89) Missing = 6
Prenatal Smoking (ever)	No: 272/72.5% Yes: 82/21.9% Missing: 21/5.6%	No: 51/57.3% Yes: 20/33.7% Missing: 8/9.0%	No: 215/77.9% Yes: 25/9.1% Missing: 36/13.0%	No: 66/78.6% Yes: 8/9.5% Missing: 10/11.9%
Parity	0: 168/44.8% 1+: 205/54.7% Missing: 2/0.5%	0: 35/39.3% 1+: 54/60.7%	0: 118/42.8% 1+: 157/56.9% Missing: 1/0.4%	0: 25/29.8% 1+: 58/69.0% Missing: 1/1.2%
Marital status	Single: 94/25.1% Married: 258/68.8% Separated: 5/1.3% Divorced: 18/4.8%	Single: 20/22.5% Married: 55/61.8% Separated: 6/6.7% Divorced: 8/9.0%	Single: 195/70.7% Married: 59/21.4% Separated: 12/4.3% Divorced: 7/2.5% Missing: 3/1.1%	Single: 53/63.1% Married: 22/26.2% Separated: 5/6.0% Divorced: 3/3.6% Missing: 1/1.2%
Poverty	310.2 (251.2) Missing = 32	268.1 (227.2) Missing = 14	134.5 (110.0) Missing = 58	121.8 (89.3) Missing = 13
Douching	Douching: 77/20.5% No douching: 173/46.1% Missing: 125/33.3%	Douching: 25/28.1% No douching: 26/29.2% Missing: 38/42.7%	Douching: 86/31.1% No douching: 68/24.6% Missing: 122/44.2%	Douching:24/28.6% No douching: 10/11.9% Missing: 50/59.5%
Depression	15.5 (11.3) Missing = 68	15.6 (10.1) Missing = 28	18.8 (10.4) Missing = 99	19.6 (11.3) Missing = 45
Negative Life Events	-6.4 (7.0) Missing =66	-8.1 (7.6) Missing =28	-7.1 (6.7) Missing = 107	-11.1 (11.8) Missing = 47

Figures

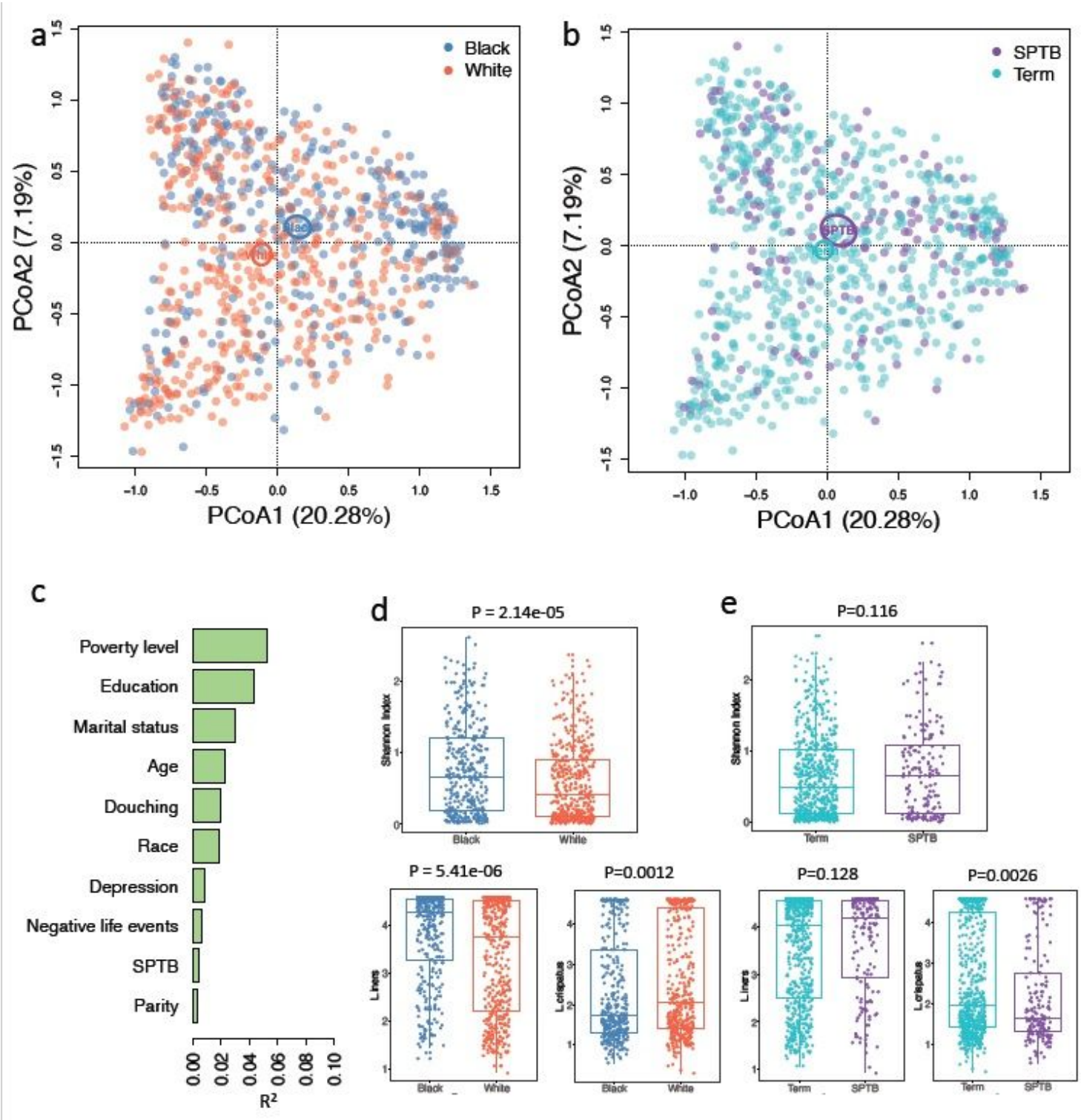


Figure 1

PCoA ordination of vaginal microbiome colored by race (a) and term/spontaneous preterm birth (SPTB) (b). The ellipse indicates a 95% confidence limit. (c) The effect sizes (R2) of host factors with PERMANOVA tests. (d) The associations between race and Shannon diversity, L. iners and L. crispatus. (e) The associations between sPTB and Shannon diversity, L. iners and L. crispatus. Significance was determined by non-parametric Wilcoxon test.

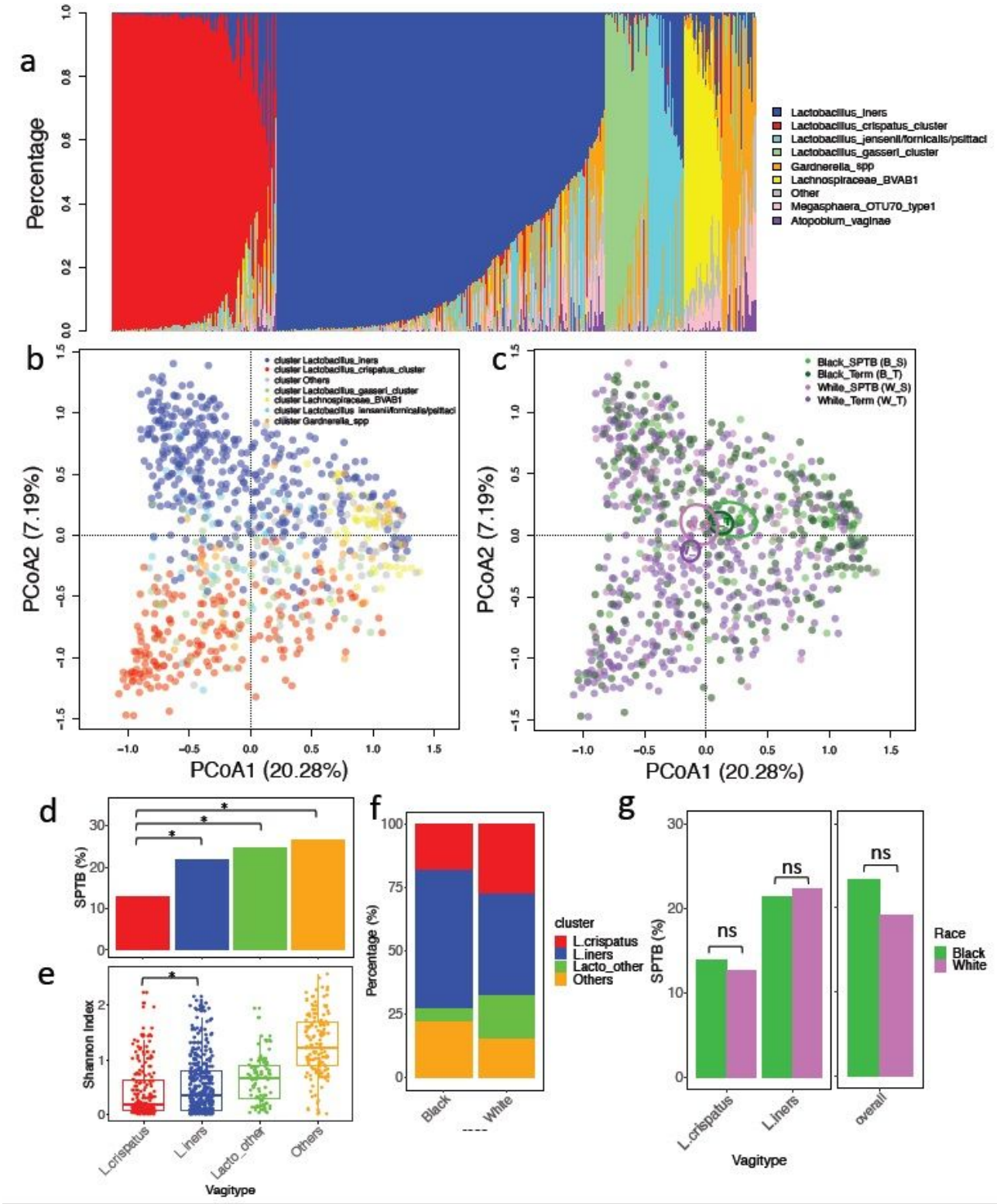
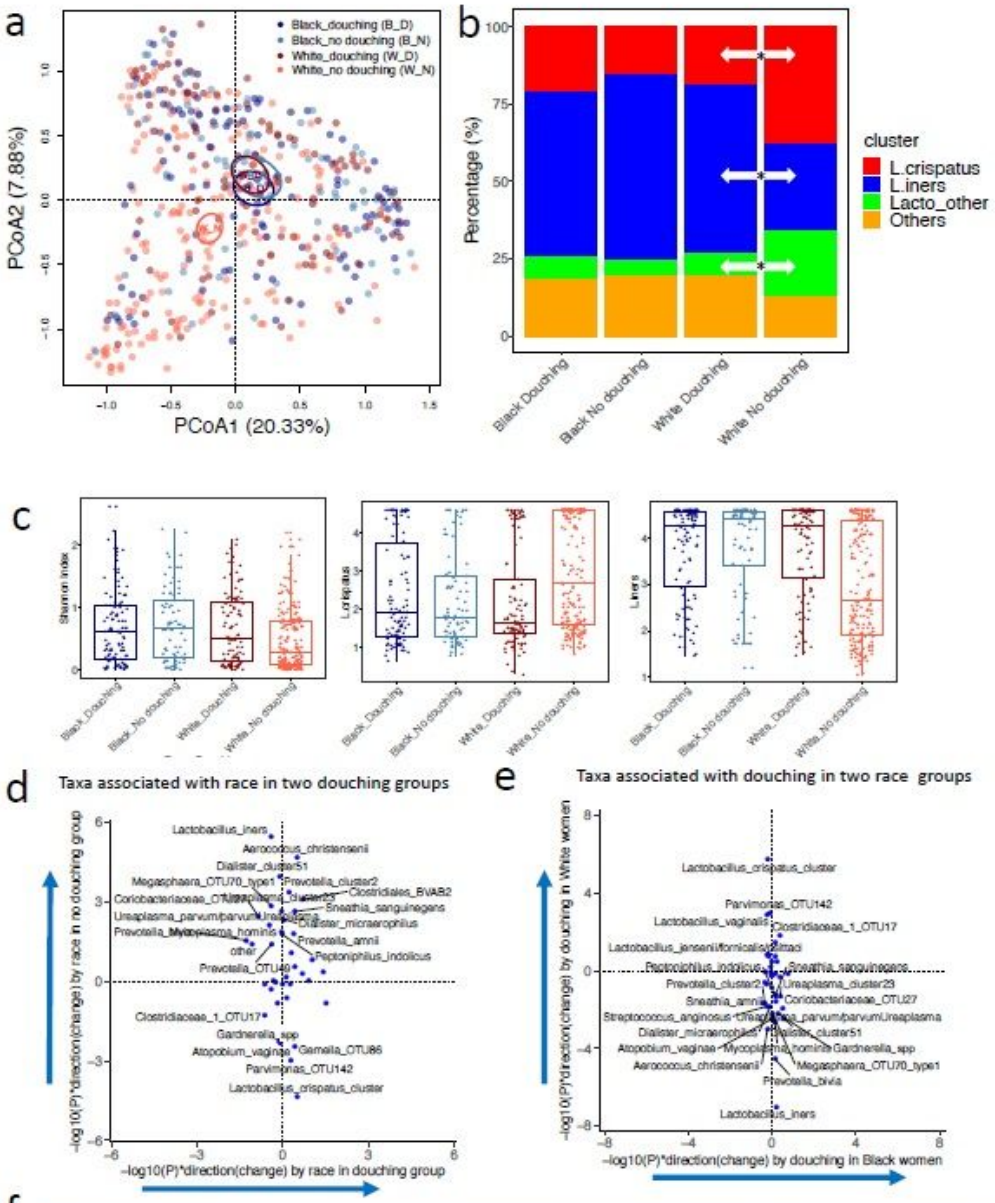


Figure 2

The vagitypes and their associations with preterm birth. (a) The vaginal microbiome of study participants showed discrete taxonomic composition that can be classified to vagitypes. (b) The PCoA ordination of the vaginal microbiome colored by vagitypes. (c) The PCoA ordination of the vaginal microbiome colored by race and term/spontaneous preterm birth (sPTB). (d) The percentage of sPTB in each of the vagitypes. Significance was determined with Fisher's exact test. (d) Shannon diversity of the vagitypes. The

Shannon diversity of L. crispatus vagitype was significantly lower than that of L. iners vagitype, estimated with Wilcoxon test. (f) Vagitype composition of Black and White women. (g) The percentage of sPTB associated with L. crispatus and L. iners vagitypes in Black and White women. Significance was determined with Fisher's exact test.



f

Association between microbiome and sPTB within categories of race and douching by PERMANOVA test

Race	Douching	#Participants	SPTB percentage	R^2	P
Black	Yes	110	21.8%	0.0129	0.137
	No	78	12.8%	0.0293	0.03 *
White	Yes	102	24.3%	0.0103	0.387
	No	199	13.0%	0.0099	0.034 *

Figure 3

The microbial variation by race and douching and its interference with microbial signatures associated with spontaneous preterm birth (sPTB). (a) PCoA ordination colored by race and douching. (b) Vagitype composition of douching and non-douching Black and White women. Significance was determined with Fisher's exact test. (c) Shannon diversity, L. crispatus and L. iners abundance by race and douching groups. The Shannon diversity, L. crispatus and L. iners abundance were significantly different for no douching White women compared to the other three groups, estimated with Wilcoxon test. (d) Comparison of microbial species associated with race in douching and no douching participants. The x-axis is -log10(P) of the associations between each taxon and race with Wilcoxon test in douching group, multiplied by the direction of changes, while the y-axis is in no douching group. The direction of arrows is towards higher abundance in Black women compared to White women. Only species with FDR<0.1 were labeled. (e) Comparison of microbial species associated with douching in Black and White participants. The x-axis is -log10(P) of the associations between each taxon and douching with Wilcoxon test in Black women, multiplied by the direction of changes, while the y-axis is in White women. The direction of arrows is towards higher abundance in the no douching group compared to the douching group. Only species with FDR<0.1 were labeled. (f) PERMANOVA tests of the associations between microbiome and term/SPTB birth in each of the race and douching groups.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SI.pdf