Selected Vaginal Bacteria and Risk of Preterm Birth: An Ecological Perspective

Ai Wen,1 Usha Srinivasan,1 Deborah Goldberg,2 John Owen,2 Carl F. Marrs,1 Dawn Misra,4 Deborah A. Wing,5 Sreelatha Ponnaluri,1 Arianna Miles-Jay,1 Brigette Bucholz,1 Khadija Abbas,1 and Betsy Foxman1

1Department of Epidemiology, and 2Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan; 3Department of Obstetrics and Gynecology, University of Alabama at Birmingham, 1700 6th Ave. South, Birmingham, Alabama 35233; 4Department of Family Medicine and Public Health Sciences, School of Medicine, Wayne State University, 3939 Woodward, Detroit, Michigan; 5Obstetrics and Gynecology, School of Medicine, University of California Irvine, 101 City Drive South, Bldg 56, Ste 800, Orange, California

We examined the community ecology of vaginal microbial samples taken from pregnant women with previous preterm birth experience to investigate whether targeted pathogenic and commensal bacteria are related to risk of preterm birth in the current pregnancy. We found a significant correlation between the community structure of selected bacteria and birth outcome, but the correlation differed among self-reported racial/ethnic groups.

Using a community ordination analysis, we observed infrequent co-occurrence of *Mycoplasma* and bacteria vaginosis associated bacteria 3 (BVAB3) among black and Hispanic participants. In addition, we found that the vaginal bacteria responded differently in different racial/ethnic groups to modifications of maternal behavioral (ie, douching and smoking) and biological traits (ie, body mass index [BMI]). Even after accounting for these maternal behaviors and traits, the selected vaginal bacteria was significantly associated with preterm birth among black and Hispanic participants. By contrast, white participants did not exhibit significant correlation between microbial community and birth outcome. Findings from this study affirm the necessity of considering women’s race/ethnicity when evaluating the correlation between vaginal bacteria and preterm birth. The study also illustrates the importance of studying the vaginal microbiota from an ecological perspective, and demonstrates the power of ecological community analysis to improve understanding of infectious disease.

**Keywords.** preterm birth; vaginal; microbial community; community analysis.

Accurately predicting and preventing preterm birth remains a challenge to the medical community: nearly 12% of all deliveries occur at <37 weeks [1]. An estimated 25%-40% of preterm births (PTB) are attributed to intrauterine infection [2, 3]. However, the relationship between the presence of pathogens and spontaneous preterm birth is complex. Despite a 2–5 fold increase in risk of PTB among women with bacterial vaginosis (BV) [4], treatment of BV using antibiotics has yielded conflicting results [4–6]. The reasons for the failure of antibiotic treatment are unclear [7]. One hypothesis is that BV diagnosis via the Nugent or Amsel criteria does not capture the range and complexity of vaginal dysbiosis that leads to PTB. This complexity can include the alternation of enzyme activities [8] and prostaglandin level [9]. Results of one prospective study support this contention, as adding additional microscopic patterns to the Nugent score improved the prediction of PTB among those with abnormal Gram stain [10]. Recently, the detection of BV using quantitative polymerase chain reaction (qPCR) has been proposed, but selection of bacteria for inclusion has not related to PTB risk [11].

New molecular methods have enabled the identification of novel BV-associated bacteria [12, 13] in addition to species traditionally identified by vaginal-fluid Gram stains [14], but the roles of these bacteria in preterm birth remain unclear. Specifically, the association of the microbial community structure to preterm birth and...
the relationship of specific pathogen occurrences with other pathogenic or commensal bacteria have yet to be described. Here we address this gap by evaluating the roles of selected bacterial taxa in preterm birth, taking account of the interactions among facultative pathogenic and commensal bacteria by using community ecology methods.

We recognize the hazard of studying the association between the maternal microbial community and preterm birth in isolation, distinct from other known predictors of preterm birth linked to maternal traits (Figure 1). These factors include black race [3], smoking [15], douching [16–19], extremes of BMI [3, 20], and shortened cervical length during pregnancy [3, 21]. The mechanisms by which these factors increase the risk of preterm birth are poorly understood; the effects may be direct, indirect, or both and possibly involve the vaginal microbiota. Direct mechanisms (Figure 1, dotted line) include vasoconstriction from tobacco smoke exposure, which reduces blood flow to the placenta [3], and factors associated with shortened cervical length, a powerful predictor of preterm birth [21]. Indirectly, these factors may have their effects through modification of the microbial community, for example, through facilitating pathogen invasion or changing the relative abundance of innate bacteria (Figure 1). Without considering how these factors modify the vaginal flora, we cannot correctly estimate the association between vaginal microbial community and preterm birth.

We conducted our study using results from a cohort of women who had experienced at least one previous spontaneous preterm birth and hence had high risk of recurrence [22]. We targeted commensal bacteria known to occur naturally in the vaginal cavity (lactic acid bacteria and Lactobacillus spp.; lactic acid bacteria include Lactobacillus, Weisella, Pediococcus, and Leuconostoc) and facultative pathogenic bacteria that have been associated with BV [4], including Atopobium spp., Clostridia-like bacterial vaginosis-associated bacterium (BVAB) 1, 2, and 3, Escherichia coli, Gardnerella vaginalis, Group B Streptococcus, Mobiluncus spp., Mycoplasma spp., and Ureaplasma spp. We compared the prevalence of these vaginal inhabitants in the microbial community during the second trimester of pregnancy.

A number of studies have shown that the vaginal microbiota differ among racial/ethnic groups [23, 24]. We used the participants’ self-report to define racial/ethnic groups and stratified the microbial analyses with these groups. Our use of the term “racial/ethnic” group includes a wide array of potential factors that may affect preterm birth, including sociological and economic factors. Using correspondence analysis (CA), a widely applied community analysis tool that can depict organismal community structure using the presence/absence of organisms [25], we first examined whether the vaginal microbial community was correlated with preterm birth for each racial/ethnic group. As shown by many ecological studies, presence/absence data can successfully produce comparable results as analyses with abundance data [26–28], especially in data with a large number of zeros as shown in our study. We then applied constrained community analysis methods that take into account the targeted risk factors (smoking, douching, and BMI) while examining the underlying relationship between the microbial community and preterm birth.

MATERIALS AND METHODS

Study Population

We tested previously collected (Gram stains) for the presence of selected bacterial genera and species collected as part of the National Institute of Child Health and Human Development-sponsored randomized trial of cerclage for PTB prevention [22]. The study protocol was approved by Human Subjects Committee at the University of Alabama at Birmingham (X991227014) and at all participating sites. The use of de-identified data from this study was deemed exempt and unregulated by the Human Subjects Committee at the University of Michigan because the original consent form provided from the collection and analysis of biological specimens as it related to studying the etiology of preterm birth.

Pregnant women carrying a singleton gestation and who had at least one prior spontaneous preterm birth between 17 and 33 weeks' gestation were enrolled. Throughout the study, if a participant had shortened cervical length (<25 mm) at one of the visits (determined by a serial transvaginal ultrasound), the participant was randomized to a cerclage group vs control group. This report only considers samples from participants who had a cervical length of at least 25 mm at all sonographic assessments (see Sample selection for community analysis for the
rationale to exclude participants with shortened cervical length. We also collected information on BMI, smoking history during pregnancy, douching before pregnancy, and the self-reported racial/ethnic group of the participant during the initial visit. To depict the effect of douching on vaginal microbial community, we excluded the small number of participants who continued douching during pregnancy from data analysis.

Sample Collection and Quantitative PCR Assay for Targeted Bacteria Taxa
Before the initial sonographic cervical length evaluation, which was scheduled between 16 weeks 0 days to 21 weeks 6 days gestation, a sterile speculum examination was performed to collect vaginal fluid from the upper one-third of the vaginal sidewalls for pH and Gram stain.

Bacterial DNA was extracted from the Gram stain slides as described elsewhere [29]. After preamplification with 8F-1492R, universal bacterial primers based on 16S rDNA, total bacterial and species specific bacterial copy numbers were identified using total bacteria primers and primers specific to each bacteria taxa (12 bacteria taxa including *Atopobium spp.*, *Clostridia*-like bacterial vaginosis-associated bacterium (BVAB)1, 2, and 3, *E. coli*, *Gardnerella vaginalis*, Group B *Streptococcus*, lactic acid bacteria, *Lactobacillus spp.*, *Mobiluncus spp.*, *Mycoplasma spp.*, and *Ureaplasma spp.*). The primer sequences of these bacteria are shown in Supplementary Table 1. Samples with copy numbers <100 were under the limit of detection and were considered to be negative. Because the DNA was extracted from previous stored slides that reasonably differed in sample volume, the total bacteria load among slides could not be considered a constant. Therefore, we used the presence/absence of bacteria taxa to construct the bacteria community of each sample.

Sample Selection for Community Analysis
To analyze the correlation between microbial community structure and preterm birth, we selected participants whose preterm birth was spontaneous and not indicated for other maternal/fetal complications. A previous study using this cohort demonstrated that the placement of cervical cerclage reduced the risk of preterm birth in women with shortened cervix <25 mm [22], which suggested that preterm birth among women with shortened cervical length may be due to “inherent mechanical weakness” [30] of the cervix. Therefore, we excluded participants who exhibited cervical shortening (<25 mm). Previous studies also showed that smoking during pregnancy may cause vaso-constriction that reduces utero-placental blood flow [3] and is associated with an increased risk of preterm birth [31]. However, smoking has also been associated with variation in vaginal bacteria composition [32, 33], which can be an alternative mechanism to explain the high risk of preterm birth among smokers. We therefore included smokers and nonsmokers in the community analysis to assess the effect of smoking on preterm birth via altered microbial community; in addition, we conducted post hoc tests on nonsmokers and smokers, respectively, to specifically evaluate the effect of smoking on preterm birth (see Supplementary Materials).

Statistical Analysis
To examine the relationship between vaginal microbial community and the birth outcome of each racial/ethnic group, we applied CA to the data on presence/absence of microbial taxa (Figure 1), stratified by the participants’ race/ethnicity. This analysis reduces the dimensions of the community data by taking into account the correlations among taxa and maximizing the correspondence between samples and microbial taxa [25]. Reducing dimensionality is a key feature of multivariate analyses, which allows further analyzing and visualizing community data in a simplified space. We applied a post-ordination permutation analysis to examine whether the samples from preterm and term participants differed in their locations in the reduced-dimension space, which would indicate that the taxon composition of the microbial community differs between preterm and term participants.

Second, we applied canonical correspondence analysis (CCA) on the microbial community data [25] stratified by the race/ethnicity of participants. This method applies constraining factors (ie, douching before pregnancy, smoking during pregnancy, and BMI) to the ordination of community data and allows post hoc examination on whether these factors contribute significantly to the composition of community. Compared to the unconstrained CA, CCA specifically extracts the variance in community data that can be explained by the factors of interest. In other words, if CCA shows insignificant correlations between the microbial community and the tested factors, it indicates that the microbial community is probably associated with factors not included in the analysis.

We conducted CCA in 2 steps. First, we tested whether any of the constraining factors significantly modify the microbial community, without including the birth outcome as a constraint to the ordination (CCA I, Figure 1). By doing this we tested whether the constraining factors modify the microbial community independent of the birth outcome. We then conducted the full CCA including both the set of constraining factors and birth outcome (CCA II, Figure 1). This analysis determines whether the structure of the microbial community is associated with birth outcome, after taking into account effect of the constraining factors (ie, douching before pregnancy, smoking during pregnancy, and BMI). A Monte Carlo permutation test was used to assess the effect of each factor included in both CCAs, while subtracting the effect caused by the other constraining factors. Such permutation-based methods have been applied widely to assess whether observed environmental factors are associated with the ordination of microbial communities.
communities [25]. We compared the CCA test to the result of CA analysis to determine whether variance in the microbial community corresponds to the preterm or term birth outcome after removing the effects of these other constraining factors. The null hypothesis for the permutation test is that the observed association between the ordination and constraining factors happens by chance. If the permutation test shows that the association happens in <5% of the hundreds of permutations (P < .05), it indicates that the observed pattern is unlikely to happen by chance. All of the community analyses were carried using the vegan package [34] in the software program R 2.15.2 [35].

RESULTS

Among the 608 vaginal Gram stain slides received, 234 were excluded from the microbiological study for the following reasons: Gram stained sample could not be amplified (n = 14); slides were associated with incomplete hospital visiting information (n = 31); and samples showed measurements that fell >5 standard deviations from the mean and were considered as outliers (n = 18). In addition, participants who experienced cervical shortening (<25 mm, n = 129) were excluded from this study (see Sample selection for community analysis for justification). Women with an inconclusive ultrasound cervical measurement were, however, included because although their cervix could not be measured precisely due to a poorly developed lower segment, the cervical length was longer than 25 mm. We also excluded the few participants who reported douching during pregnancy (n = 12) because the exposure to douching was quite variable in when it was ascertained during pregnancy (from 17 to 33 weeks representing a very heterogeneous exposure), and the number of samples was too low to be included as a factor in the community analysis. Finally, we excluded women who reported their race/ethnicity as “Other” (n = 30); all of these participants delivered preterm (ie, gestation age <37 weeks), effectively leaving no comparison group for analyses of preterm birth risk in this small subgroup. The characteristics of the selected sample are shown in Table 1.

Without applying the constraining factors associated with the maternal traits, the microbial community ordinated by CA was significantly associated with birth status (preterm vs term birth) within the black and Hispanic groups (Table 2, CA; Figure 2A and 2E). In addition, the correlation remains significant even after constraining by known preterm birth risk factors (Table 2, CCA II; Figure 2B and 2F). Among the white participants (Table 2), we did not observe significant correlation between the microbial community and birth status with or without the constraining factors (Figure 2C and 2D).

Black Women

Among the black participants, BMI was significantly correlated with the microbial community, whereas other factors considered were not related to the community structure (Table 2, CCA I). After subtracting the effect of all the constraining factors, the correlation (P = .005) between microbial community and birth outcome persisted (Table 2, CCA II; Figure 2B). Within the black subgroup, we further investigated the interaction and cooccurrence of particular bacteria taxa and their association with preterm birth. Specifically, the ordination analysis suggested the presence of BVAB3 was associated with reduced risk of PTB, whereas the presence of Mycoplasma was associated with higher risk of PTB, even after subtracting the effect of other constraining factors (Figure 2B). An odds ratio (OR) analysis showed that the presence of Mycoplasma significantly increases the chance of preterm birth (OR = 5.70 [2.40, 14.4], P < .001) whereas BVAB3 drastically decreases the risk of preterm birth (OR = 0.13 [0.036, 0.38], P ≤ .001). The ordination result also showed that these 2 bacteria taxa locate at opposite extremes along the first CCA axis, which, because the ordination was conducted using presence/absence data, indicates their rare cooccurrence. In fact, these 2 bacteria taxa only co-occur in 7 of the 59 participants who had at least one of the taxa present (Figure 3A).

Hispanic Women

Among Hispanic participants, the correlation between microbial community and birth outcome also remained significant after removing the effects of constraining factors (Table 2 CCA II, Figure 2F). Results of the CCA II showed that Mycoplasma is still strongly correlated with preterm birth (Figure 2F, OR = 4.45 [1.69, 11.97], P < .01), but the negative association with BVAB3 was only marginally significant, although the association

Table 1. Demographic, Biological, and Behavioral Characteristics of 374 Participants Included in Ecological Analysis of Selected Vaginal Bacteria and Preterm Birth

<table>
<thead>
<tr>
<th></th>
<th>Black (N = 127)</th>
<th>Hispanic (N = 141)</th>
<th>White (N = 106)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preterm</td>
<td>55 (43.3%)</td>
<td>39 (27.7%)</td>
<td>32 (30.2%)</td>
</tr>
<tr>
<td>Term</td>
<td>72 (66.7%)</td>
<td>102 (72.3%)</td>
<td>74 (69.8%)</td>
</tr>
<tr>
<td>Age (mean ± SD)</td>
<td>26.0 ± 4.7</td>
<td>27.5 ± 5.7</td>
<td>27.5 ± 5.9</td>
</tr>
<tr>
<td>BMI (mean ± SD)</td>
<td>31.3 ± 8.6</td>
<td>27.6 ± 5.2</td>
<td>27.3 ± 6.4</td>
</tr>
<tr>
<td>Smoking during pregnancy</td>
<td>22 (17.3%)</td>
<td>2 (1.4%)</td>
<td>26 (24.5%)</td>
</tr>
<tr>
<td>Douching before pregnancy</td>
<td>35(27.6%)</td>
<td>23 (16.3%)</td>
<td>15 (14.2%)</td>
</tr>
<tr>
<td>Nugent score ≤3</td>
<td>61 (48.0%)</td>
<td>91 (64.5%)</td>
<td>78 (73.6%)</td>
</tr>
<tr>
<td>Nugent score between 4 and 6</td>
<td>44 (34.6%)</td>
<td>38 (26.9%)</td>
<td>24 (22.6%)</td>
</tr>
<tr>
<td>Nugent score ≥7</td>
<td>22 (17.3%)</td>
<td>12 (8.5%)</td>
<td>4 (3.8%)</td>
</tr>
</tbody>
</table>

Participants were at high risk of preterm birth; in all cervical length remained ≥25 mm during the study period.

Abbreviations: BMI, body mass index; SD, standard deviation.
was in the same direction (OR = 0.19 [0.0076,1.01], P = .068).
Similar to black participants, BVAB3 and *Mycoplasma* remain at the 2 extremes along the first CCA axis, indicating that these 2 bacteria have a very low incidence of co-occurrence (Figure 3B).

**DISCUSSION**

Among 374 gravidas who had at least one previous spontaneous preterm birth and who maintained a cervical length of ≥25 mm in this index pregnancy, the vaginal bacterial community in the second trimester was correlated with birth outcome, but the correlation depended on the race/ethnicity of the mother. Among black and Hispanic women, even after controlling for selected maternal behavioral and biological characteristics, there remained residual variance in the bacterial community that significantly contributed to preterm birth. However, among white women, the correlation between microbial community and preterm birth was not significant with or without accounting for the maternal behavioral and physiologic characteristics.

Results from our study yield one possible explanation to the widely observed disparity of preterm birth risk among self-reported racial/ethnic groups [3]. Deep sequencing of the
vaginal microbial community has revealed significant differences in community structure among healthy women from different racial/ethnic groups [23]. These different structures may explain the variable effects of some of the bacterial taxa that are known to be BV related included in our study. *Mycoplasma* was positively associated with preterm birth among both black and Hispanic participants (Figure 2A, 2B, 2E and 2F) but had no significant effect among white participants in our sample. A companion study using the same study cohort showed that the occurrence of BV associated bacteria does not exhibit significant differences among the 3 racial/ethnic groups [36]. This suggested that the nonsignificant association between *Mycoplasma* and PTB among white participants is not caused by lower prevalence of this bacterium. We also found that the presence of BVAB3 had highly significant negative correlation with PTB among black participants. This effect was not significant among Hispanic participants but did show the same trend (Figure 2E and 2F). Consistent with this, the community analyses also showed these 2 bacteria taxa tended not to co-occur among black and Hispanic participants (Figure 3). The role of *Mycoplasma* in preterm birth has been suggested by many previous studies, but the exact etiology remains unclear [37]. Our study suggests that the inconsistent associations of *Mycoplasma* with preterm birth reported may be due to failure to take into account its interaction with other taxa (such as BVAB3) and maternal race/ethnicity.

The discovery of an apparent negative association between BVAB3 and PTB was unexpected and needs further evaluation. This species was only recently described using molecular tools [12] and has been described as a specific indicator of BV. Its effect in preterm birth had not been evaluated in published literature previous to this study. Although BV has been positively associated with preterm birth, its complex, multifactorial etiology has confounded the link between BV-associated bacteria and preterm birth. More targeted studies on these bacteria and preterm birth are essential to build a comprehensive understanding toward the microbiological etiology of BV and preterm birth.

Many maternal behavioral and physiological traits modify the rates of preterm birth [3], but whether these host-specific characteristics affect preterm birth via the mediation of vaginal bacterial community is unclear. Although our study was not designed to directly test the effects of these factors on vaginal microbiology, the ordination technique we applied allowed us to assess the marginal effect of these factors on the microbial community while controlling for other factors. We found that among the black participants BMI was significantly correlated to the structure of the microbial community (Table 2, Figure 2B). As suggested by Figure 2B, further analysis using ANOVA indicated that BMI is negatively correlated with the presence of *Mycoplasma* (mean BMI difference = −4.25, *P* = .018) and BVAB2 (mean BMI difference = −4.9, *P* = .004) in this racial group. However, after taking account of this effect, the bacteria community was still shown to be correlated with birth status (Table 2). Similarly, among the Hispanic participants, the vaginal bacteria community is correlated with preterm birth after taking account of the constraining factors. In this race/ethnicity group, douching before pregnancy shifted the bacterial community, but the effect is only marginally significant (Table 2, *P* = .074). However, further examination showed that douching before pregnancy significantly reduced the prevalence of BVAB1 (OR = 0.26 [0.056, 0.95], *P* = .026) and *Gardnerella vaginalis* (OR = 0.21 [0.029, 0.79], *P* = .021) in this race/ethnicity group. A previous study showed that douching 6 months before pregnancy reduces the risk of preterm birth among black women with low income [18]. In our study, the reduced prevalence of BVAB1 and *Gardnerella vaginalis* with douching was an indicator for shifts in vaginal bacterial community that may change the risk for preterm birth. Note that in this study, the percentage of participants who douch before pregnancy (Table 1) is low compared to previous similar studies [17, 18]. It may be the result of different participant selection criteria such as socioeconomic status, or exclusion of participants with shortened cervical length, etc.

We excluded participants with shortened cervical length from the microbiological analysis because of the well-established
link between cervical shortening and preterm birth via non-microbial pathways [3, 22, 30]. Nevertheless, it is possible that the cervical shortening itself involves interactions between the host’s vagina and changes in host’s vaginal microbiome. More detailed analysis on the correlation between targeted bacteria genus, the shortening of cervical length, and preterm birth is detailed analysis on the correlation between targeted bacteria and preterm birth; future studies should focus on the major bacterial taxa present in the vaginal samples.

We conclude that the vaginal microbial community at the second trimester can play an important role in determining the risk of preterm birth among women with previous preterm birth experience. This study, consistent with recent studies on racial/ethnic groups and bacterial communities, reiterates that including diverse populations of pregnant women is critical for understanding the etiology of infection-associated preterm birth, because the microbial community exhibited strong structural differences among racial/ethnic groups. Our study also suggests that the interaction among bacteria taxa may critically affect the bacteria’s role in preterm birth and hence demonstrates the potential analytical value of community ecology methods in infectious disease studies.

### Supplementary Data

Supplementary materials are available at The Journal of Infectious Diseases online (http://jid.oxfordjournals.org/). Supplementary materials consist of data provided by the author that are published to benefit the reader. The posted materials are not copyedited. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

### Notes

Acknowledgments. We thank Dr Jeff Szychowski for data management.

Financial support. This work was supported by the National Institutes of Health [grant R01 HD038098 to B. F. and D. M.].

Potential conflict of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

### References