BJOG An International Journal of Obstetrics and Gynaecology

DOI: 10.1111/1471-0528.15299
www.bjog.org

Epidemiology

# Vaginal microbiome in early pregnancy and subsequent risk of spontaneous preterm birth: a case–control study

N Tabatabaei,[a,b] AM Eren,[c] LB Barreiro,[a,d] V Yotova,[a,d] A Dumaine,[a,d] C Allard,[e] WD Fraser[a,b,e]

[a] Research Centre of CHU Sainte-Justine, University of Montreal, Montreal, QC, Canada [b] Department of Obstetrics and Gynaecology, University of Sherbrooke, Sherbrooke, QC, Canada [c] Department of Medicine, University of Chicago, Chicago, IL, USA [d] Department of Paediatrics, Research Center of CHU Sainte-Justine, University of Montreal, Montreal, QC, Canada [e] Research Centre of Sherbrooke University Hospital Centre (CRCHUS), Sherbrooke, QC, Canada

*Correspondence:* WD Fraser, Department of Obstetrics and Gynaecology, University of Sherbrooke, 3001, 12e Avenue Nord Sherbrooke, QC J1H 5N4, Canada. Email: william.fraser@usherbrooke.ca

*Accepted 23 April 2018. Published Online 27 June 2018.*

**Objectives** To explore differences in the vaginal microbiome between preterm and term deliveries.

**Design** Nested case–control study in 3D cohort (design, develop, discover).

**Setting** Quebec, Canada.

**Sample** Ninety-four women with spontaneous preterm birth as cases [17 early (<34 weeks) and 77 late (34–36 weeks) preterm birth] and 356 women as controls with term delivery (≥37 weeks).

**Methods** To assess the vaginal microbiome by sequencing the V4 region of the 16S ribosomal RNA (rRNA) gene in swabs self-collected during early pregnancy.

**Main outcome measures** Comparison of relative abundance of bacterial operational taxonomic units and oligotypes and identifying vaginal community state types (CSTs) in early or late spontaneous preterm and term deliveries.

**Results** *Lactobacillus gasseri/ Lactobacillus johnsonii* (coefficient −5.36, 95% CI −8.07 to −2.65), *Lactobacillus crispatus* (99%)/ *Lactobacillus acidophilus* (99%) (−4.58, 95% CI −6.20 to −2.96), *Lactobacillus iners* (99%)/ *Ralstonia solanacearum* (99%) (−3.98, 95% CI −6.48 to −1.47) and *Bifidobacterium longum*/ *Bifidobacterium breve* (−8.84, 95% CI −12.96 to −4.73) were associated with decreased risk of early but not late preterm birth. Six vaginal CSTs were identified: four dominated by *Lactobacillus*; one with presence of bacterial vaginosis-associated bacteria (*Gardnerella vaginalis*, *Atopobium vaginae* and *Veillonellaceae* bacterium) (CST IV); and one with nondominance of *Lactobacillus* (CST VI). CST IV was associated with increased risk of early (4.22, 95% CI 1.24–24.85) but not late (1.63, 95% CI 0.68–5.04) preterm birth, compared with CST VI.

**Conclusions** *Lactobacillus gasseri*/*L. johnsonii*, *L. crispatus*/ *L. acidophilus*, *L. iners*/*R. solanacearum* and *B. longum*/*B. breve* may be associated with decreased risk of early preterm birth. A bacterial vaginosis-related vaginal CST versus a CST nondominated by *Lactobacillus* may be associated with increased risk of early preterm birth.

**Keywords** 16S rRNA, bacterial vaginosis, *Lactobacillus*, preterm birth, vaginal microbiome.

**Tweetable abstract** Largest study of its kind finds certain species of vaginal *Lactobacillus* + *Bifidobacterium* may relate to lower risk of preterm birth.

**Linked article** This article is commented on by SS Witkin, p. 359 in this issue. To view this mini commentary visit https://doi.org/10.1111/1471-0528.15300.

*Please cite this paper as:* Tabatabaei N, Eren AM, Barreiro LB, Yotova V, Dumaine A, Allard C, Fraser WD. Vaginal microbiome in early pregnancy and subsequent risk of spontaneous preterm birth: a case–control study. BJOG 2019;126:349–358.

## Introduction

Bacterial vaginosis (BV) is a state of altered vaginal microbiome that has been associated with increased risk of preterm birth, especially early in pregnancy.[1–3] Bacterial vaginosis is diagnosed by Nugent score[4] but it is subjective and requires specialised training.[5] Recently, sequencing of the 16S ribosomal RNA (rRNA) gene has resulted in precise bacterial identification in different tissues including the vagina.[6,7]

The association between maternal vaginal microbiome and risk of preterm birth is controversial.[2,8–11] In a US case–control study, the risk of preterm birth was higher in patients within a community state type (CST) poor in

*Lactobacillus* but with highly abundant *Gardnerella* or *Ureaplasma*.[2] In a US cohort,[8] no correlation was observed between low abundance or absence of *Lactobacillus* and the risk of preterm birth, despite lower intracommunity diversity [Shannon diversity index (SDI)] in preterm versus term white population.[8] A UK study reported that a dominance of *Lactobacillus crispatus* in the vaginal microbiota at 16 weeks of gestation is protective against preterm birth (<34 weeks), whereas *Lactobacillus iners* is a risk factor for preterm birth in women at high risk of preterm birth.[11] This suggests the importance of identification of *Lactobacillus* to species level. They recently reported a similar protective effect of *L. crispatus* on preterm birth but the association between lower *Lactobacillus*, higher *Gardnerella* and preterm birth was only seen among white population and African Americans.[12] Additionally, a cross-sectional US study (mainly African Americans) reported no distinct taxa across pregnancy in association with preterm birth.[10] Similarly, no differences in SDI, bacterial taxa and vaginal CSTs were reported between women with early preterm (<34 weeks) versus term deliveries, despite mainly being African Americans,[9] which is an ethnicity associated with higher risk of both BV and preterm birth.[6,13] These studies suggest that the association between the vaginal microbiome and preterm birth is population-dependent.[12]

Previous studies on the vaginal microbiome and preterm birth had small sample size,[2,8–10] inconsistency in collection of vaginal swabs across pregnancy[8,10] and limited information on spontaneous preterm births.[2,14] Therefore, the objective of the current large case–control study (mainly white European population) was to compare the composition of the bacterial community in vaginal swabs collected consistently in early pregnancy of women with early (<34 weeks) and late (34–36 weeks) spontaneous preterm and term (≥37 weeks) deliveries and to identify vaginal CSTs in association with risk of early and late preterm versus term birth.

# Methods

## Sample collection and study design

This is a case–control study nested in the 3D pregnancy cohort (design, develop, discover), which included singleton pregnant women ($n$ = 2366)[15] recruited in nine Quebec hospitals during the first trimester ($8^{+0}$–$13^{+6}$ weeks) of pregnancy. More details of the 3D pregnancy data collection cohort can be found in the Supplementary material (Appendix S1A). All women provided informed consent before recruitment. From the initial cohort, there were 120 cases of preterm birth (<37 weeks of gestation) and we performed simple random selection of 360 controls who had term delivery (> 37 weeks of gestation) in their current pregnancy. To have a statistical power of 80% for the detection of ~ 3% difference in prevalence of a Nugent score > 7 (BV) at early pregnancy between preterm and term deliveries, selection of a control to case ratio of 3 : 1 is required.[16] We then excluded women with cervical cerclage in the present or any past pregnancy (due to cervical insufficiency and threatened miscarriage/preterm birth) or uterine malformation, because these women are at higher risk for preterm birth.[11,17] Patients for whom a first-trimester vaginal swab was missing were excluded ($n$ = 6 preterm and $n$ = 1 term). To exclusively explore spontaneous preterm births, patients with pre-eclampsia ($n$ = 13 preterm and $n$ = 3 term) and those who were induced for medical reasons (e.g. fetal growth restriction) before preterm birth ($n$ = 7 preterm) were also excluded. The final study sample therefore consisted of 94 cases of spontaneous preterm birth and 356 term deliveries (controls). Among the cases, there were $n$ = 17 early (<34 weeks) and $n$ = 77 late (34–36 weeks) preterm births. The consort flow diagram for the study is presented in the Supplementary material (Figure S1). This study was approved by the Ethics Committee of the Research Centre of CHU Sainte Justine.

Gestational age at delivery was assessed by maternal last menstrual period or through first-trimester ultrasound assessments of the crown–rump length.[18] We defined preterm birth as gestational age at delivery of <37 weeks and categorised preterm birth into early (<34 weeks) and late (34–36 weeks) preterm birth.

## Vaginal samples

### *Sample collection*

In the first trimester of pregnancy ($8^{+0}$–$13^{+6}$ weeks), two vaginal swabs were self-collected after instructions were provided by trained research staff. Studies have shown strong validity for the overall and morphotype-specific scores comparing self-collected swabs with those collected by trained health professionals[19] and high intra-rater and inter-rater reliabilities.[20] One vaginal swab was rolled onto glass slides, air-dried, Gram-stained and examined under oil immersion for vaginal microbiota assessment using the Nugent score.[4] The other vaginal swab was placed in a tube without any buffer and immediately stored at −80°C until assayed.

### *Nugent score*

The Nugent score was defined as: no bacterial vaginosis (0–3 score), intermediate bacterial vaginosis (4–6 score) or BV (7–10 score).[4] For more details, refer to the Supplementary material (Appendix S1B).

### *DNA extraction from vaginal swabs*

We used the procedure previously described by Ravel et al.[6] Details can be found in the Supplementary material (Appendix S1C).

*Sequencing of barcoded 16S rRNA gene amplicons*
Primers 515F and 806R were used for polymerase chain reaction amplification of the V4 hypervariable regions of 16S rRNA genes. The primers used for this study are presented in the Supplementary material (Table S1).

*Bioinformatic analyses*
We used the ILLUMINA-UTILS library v1.4.8[21] to demultiplex raw sequencing reads and merge partially overlapping paired-end reads into high-quality reads for downstream analyses. Details can be found in the Supplementary material (Appendix S1D).

*Statistical analyses*
We used count regression models to analyse read count data by assuming a Poisson or negative binomial distribution of the response[9] similar to Romero et al.[9] For the purpose of exploring differences in bacterial taxa (global alignment for sequence taxonomy) and oligotype (minimum entropy decomposition) relative abundance between early and late preterm versus term deliveries, we analysed the relative abundance of one taxa or oligotype at a time. Details can be found in the Supplementary material (Appendix S1E).

The reported estimated coefficient (coeff) represents the expected change in log relative abundance between women who developed early and late spontaneous preterm birth and those who had a term delivery. Adjustments were made by false discovery rate,[22] and the adjusted *P*-value was reported. Models were also adjusted for confounding factors in the association between vaginal microbiome and spontaneous preterm birth, which were maternal age, pre-pregnancy body mass index, ethnicity, parity and smoking history.

Continuous data were analysed by *t*-tests/analyses of variance for normally distributed data and Mann–Whitney for non-normally distributed data. Categorical variables were compared by Pearson's chi-square test or Fisher's exact test ($n < 5$), and a logistic regression (Wald method)[23] was used to determine the risk of spontaneous preterm births of CSTs compared with CST IV. Statistics were run on SAS 9.3 and R version 3.3.1.[24]

### Clustering of bacterial communities into community state types

Clustering of the bacterial communities into CSTs was performed as previously reported.[2] For more details, refer to the Supplementary material (Appendix S1F). This type of clustering effectively separated vaginal communities into six different CSTs: four that were dominated by different *Lactobacillus* species; one CST with greater diversity and presence of BV-associated bacteria (*Gardnerella vaginalis*, *Atopobium vaginae* and *Veillonellaceae* bacterium) (CST IV); and one CST with lower diversity and nondominance of *Lactobacillus* species (CST VI). Those analyses were carried out with R version 3.3.1.[24]

## Results

### Characteristics of the study population

Maternal characteristics are presented in Table 1. There were no significant differences between spontaneous preterm cases and controls in sociodemographic status, ethnicity, family income, parity, presence of vaginal infection, pre-pregnancy body mass index and smoking history.

### Vaginal microbiome and Nugent scores

We did not observe any significant correlations between relative abundance of vaginal microbial community composition and Nugent score categories (no BV, intermediate BV and BV) or for overall Nugent scores (0–10) (data not shown). Additionally, the microbial diversity (SDI) was not different between Nugent score categories ($F = 0.390$, $P = 0.677$) (see Supplementary material, Figure S2) or among overall Nugent scores (0–10) ($F = 1.753$, $P = 0.067$) (see Supplementary material, Figure S3).

### The vaginal microbial community composition during early pregnancy in women who develop early and late spontaneous preterm birth versus term delivery

The distribution of vaginal microbial relative abundance composition between pregnant women who had early (<34 weeks of gestation) and late (34–36 weeks of gestation) preterm versus term (≥37 weeks of gestation) deliveries is presented for oligotypes (Table 2) and taxonomy (see Supplementary material, Table S2). Among bacterial oligotypes, *Lactobacillus gasseri/Lactobacillus johnsonii* (coeff −5.36, 95% CI −8.07 to −2.65), *Lactobacillus crispatus* (99%)/*Lactobacillus acidophilus* (99%) (coeff −4.58, 95% CI −6.20 to −2.96) and *Lactobacillus iners* (99%)/*Ralstonia solanacearum* (99%) (coeff −3.98, 95% CI −6.48 to −1.47) were associated with decreased risk of early spontaneous preterm birth. However, these oligotypes were not significantly associated with risk of late spontaneous preterm versus term deliveries (Table 2; see Supplementary material, Figure S4A–C). *Bifidobacterium longum/Bifidobacterium breve* was also associated with decreased risk of early (coeff −8.84, 95% CI −12.96 to −4.73) but not late (coeff −0.55, 95% CI −2.90 to 1.80) spontaneous preterm birth (Table 2; see Supplementary material, Figure S4D. Other vaginal microbial oligotype relative abundance composition was not significantly different between women who had early and late preterm versus term deliveries (Table 2).

The microbial diversity for oligotypes was not different between women who had an early and late preterm delivery and those who had a term delivery (early preterm birth: SDI

**Table 1.** Demographic and clinical characteristics for spontaneous preterm birth and term deliveries

| Variables | Total population | | *P*-values |
|---|---|---|---|
| | Term, *n* (%) | Preterm, *n* (%) | |
| **Maternal age (years)** | | | |
| < 35 | 134 (37.7) | 33 (35.1) | 0.638 |
| ≥ 35 | 221 (62.3) | 61 (64.9) | |
| **Ethnicity** | | | |
| White European | 262 (73.6) | 65 (69.1) | 0.333 |
| Black African | 16 (4.5) | 10 (10.6) | |
| African American | 6 (1.7) | 0 (0) | |
| East Asian | 13 (3.7) | 2 (2.1) | |
| South Asian | 4 (1.1) | 0 (0) | |
| Arab | 24 (6.7) | 7 (7.5) | |
| South/Central American | 20 (5.6) | 6 (6.4) | |
| Canadian aboriginal | 1 (0.3) | 0 (0) | |
| Other | 10 (2.8) | 4 (4.3) | |
| **Parity** | | | |
| 0 | 168 (47.2) | 49 (52.1) | 0.603 |
| 1 | 130 (36.5) | 33 (35.1) | |
| ≥ 2 | 58 (16.3) | 12 (12.8) | |
| **Marital status** | | | |
| Married | 139 (39.0) | 36 (38.3) | 0.802 |
| Common law | 197 (55.3) | 51 (54.3) | |
| Other | 20 (5.6) | 7 (7.4) | |
| **Education** | | | |
| Some college | 45 (12.7) | 13 (14.0) | 0.916 |
| College | 103 (29.1) | 26 (28.0) | |
| Undergraduate | 139 (39.3) | 34 (36.6) | |
| Graduate | 67 (18.9) | 20 (21.5) | |
| **Household income** | | | |
| < $60 000 | 109 (31.9) | 28 (33.1) | 0.683 |
| $60–$100 000 | 121 (35.4) | 36 (40.0) | |
| ≥ $100 000 | 112 (32.7) | 26 (28.9) | |
| **Working status** | | | |
| Unemployed | 244 (68.4) | 70 (74.5) | 0.707 |
| Part-time | 39 (10.7) | 8 (8.5) | |
| Full-time | 58 (16.7) | 12 (12.8) | |
| Not in labour force | 15 (4.2) | 4 (4.3) | |
| **Presence of vaginal infection** | | | |
| Yes | 17 (4.8) | 6 (6.5) | 0.518 |
| No | 338 (95.2) | 87 (93.5) | |
| **Pre-pregnancy body mass index (kg/m$^2$)** | | | |
| < 18.5 | 21 (6.1) | 6 (6.6) | 0.141 |
| 18.5–24.9 | 224 (65.1) | 53 (58.2) | |
| 25.0–29.9 | 56 (16.3) | 12 (13.2) | |
| ≥ 30 | 43 (12.5) | 20 (22.0) | |
| **Smoking history** | | | |
| Never | 168 (47.2) | 49 (52.1) | 0.603 |
| Stopped at pregnancy | 130 (36.5) | 33 (35.1) | |
| Current smoker | 58 (16.3) | 12 (12.8) | |
| **Total** | 356 (79.1) | 94 (20.9) | – |

Distributions are compared by Pearson chi-square test.

median 0.95; interquartile range 0.39–2.33, late preterm birth: SDI median 0.99; interquartile range 0.41–1.68 and term delivery: SDI median 0.84; interquartile range 0.41–1.47, $P = 0.646$).

Additionally, no differences were observed in vaginal microbial taxonomy relative abundance composition between pregnant women who had early and late preterm versus term deliveries except for *Bifidobacterium*, which was associated with decreased risk of early preterm birth (coeff −5.29, 95% CI −7.88 to −2.70)), but this association was not significant for late preterm birth (coeff −0.64, 95% CI −2.34 to 1.07) (see Supplementary material, Table S2 and Figure S5).

### Vaginal microbial community state types at early pregnancy in women who develop early and late spontaneous preterm birth versus term deliveries

The heat map of the relative abundance of the 25 most abundant oligotypes in the vaginal communities of pregnant women by early and late preterm birth and term delivery is presented in Figure 1. Associations between vaginal CSTs and early and late spontaneous preterm birth versus term deliveries are presented in the Supplementary material (Table S3). Overall, frequencies of CST I, II, III, IV, V and VI in the entire sample were 35.6, 7.1, 25.1, 10.2, 7.1 and 14.9%, respectively. There were no differences in the overall frequency of the different CSTs between women who delivered early and late preterm and those who delivered at term (see Supplementary material, Table S3, $P = 0.1429$). However, the frequency of early spontaneous preterm birth was higher in CST IV than in CST VI (15.2 versus 3.8%, $P = 0.026$, see Supplementary material, Table S3). The microbial diversity (SDI) was significantly different between the CSTs (Figure 2, $F = 44.26$, $P < 0.0001$) and CST IV compared with CST VI was associated with increased risk of early (4.22, 95% CI 1.24–24.85) but not late (1.63, 95% CI 0.68–5.04) spontaneous preterm delivery (see Supplementary material, Table S3). We also explored the distribution of vaginal CSTs by ethnicity (see Supplementary material, Table S4). More details can be found in Appendix S2 (see Supplementary material).

## Discussion

### Main findings

This is the largest next-generation sequencing-based analysis to date with a nested case–control design exploring differences in vaginal microbiome composition between spontaneous preterm birth and term deliveries. Our results suggest that *L. gasseri*/*L. johnsonii*, *L. crispatus* (99%)/*L. acidophilus* (99%), *L. iners* (99%)/*R. solanacearum* (99%) and *B. longum*/*B. breve* may be associated with decreased risk of early but not late spontaneous preterm birth. Additionally, a vaginal CST with high diversity and presence of BV-associated bacteria (*G. vaginalis*, *A. vaginae* and *Veillonellaceae* bacterium) may be associated with

**Table 2.** Vaginal microbial oligotype differential relative abundance composition between pregnant women who had early (<34 weeks of gestation) and late (34–36 weeks of gestation) preterm versus term (≥37 weeks gestation) deliveries

| Vaginal microbial community (oligotype) | PLEM AIC | NBLEM AIC | ZINBLEM AIC | Best AIC | <34 versus ≥37 weeks | | 34–36 versus ≥37 weeks | |
|---|---|---|---|---|---|---|---|---|
| | | | | | Estimated coefficient (95% CI) | Adjusted *P*-value* | Estimated coefficient (95% CI) | Adjusted *P*-value* |
| *Lactobacillus crispatus/Lactobacillus acidophilus* | 6 311 514 | 7431.0 | NA | NBLEM | −0.26 (−1.34 to 0.82) | 0.8762 | 0.11 (−0.43 to 0.65) | 0.9309 |
| *Lactobacillus iners/Ralstonia solanacearum* | 5 705 798 | 6495 | NA | NBLEM | −0.10 (−1.34 to 1.14) | 0.9346 | 0.39 (−0.23 to 1.01) | 0.9309 |
| *Lactobacillus gasseri/Lactobacillus johnsonii* | 2 011 641 | 3294 | NA | NBLEM | −5.36 (−8.07 to −2.65) | 0.0033** | 0.20 (−0.93 to 1.33) | 0.9309 |
| *Lactobacillus jensenii/Lactobacillus* sp. S421 | 2 773 686 | 4520 | NA | NBLEM | 0.41 (−1.32 to 2.14) | 0.8762 | −0.31 (−1.15 to 0.54) | 0.9309 |
| *Lactobacillus* sp. S27 | 2 630 028 | 4367 | NA | NBLEM | −1.38 (−2.95 to 0.19) | 0.3403 | −0.44 (−1.17 to 0.30) | 0.9309 |
| *Gardnerella* sp. S494/*Gardnerella vaginalis* | 1 634 928 | 3146 | NA | NBLEM | 0.98 (−1.09 to 3.06) | 0.6923 | 0.37 (−0.78 to 1.52) | 0.9309 |
| *Bifidobacterium longum/Bifidobacterium breve* | 1 545 863 | 1650 | NA | NBLEM | −8.84 (−12.96 to −4.73) | 0.0010** | −0.55 (−2.90 to 1.80) | 0.9309 |
| *Atopobium vaginae/Atopobium* sp. S7MSR3 | 1 293 716 | 2583 | NA | NBLEM | 0.89 (−1.38 to 3.16) | 0.7819 | −0.18 (−1.67 to 1.32) | 0.9309 |
| *Lactobacillus crispatus* (99%)/*Lactobacillus acidophilus* (99%) | 1 109 517 | 3422 | NA | NBLEM | −4.58 (−6.20 to −2.96) | 0.0000** | −1.55 (−2.62 to −0.47) | 0.4367 |
| *Gardnerella vaginalis/Gardnerella* sp. S4-2 | 934 416 | 2658 | NA | NBLEM | 0.74 (−1.47 to 2.94) | 0.8237 | 0.63 (−0.66 to 1.92) | 0.9309 |
| Uncultured bacterial clone | 477 470 | 940.4 | NA | NBLEM | −2.55 (−8.49 to 3.40) | 0.7421 | −2.09 (−5.79 to 1.60) | 0.9309 |
| *Veillonellaceae* bacterium | 605 609 | 1575 | NA | NBLEM | 1.76 (−1.34 to 4.85) | 0.6501 | −1.20 (−3.40 to 0.99) | 0.9309 |
| *Streptococcus anginosus* | 278 265 | 2125 | NA | NBLEM | 0.38 (−1.67 to 2.43) | 0.9067 | 0.84 (−0.24 to 1.93) | 0.8897 |
| *Streptococcus agalactiae* | 193 441 | 771.2 | NA | NBLEM | 5.27 (0.60 to 9.94) | 0.1331 | 2.95 (0.33 to 5.58) | 0.4752 |
| *Lactobacillus iners* (99%) | 409 250 | 2423 | NA | NBLEM | −2.44 (−4.46 to −0.43) | 0.0980 | −1.25 (−2.39 to −0.10) | 0.4823 |
| *Alloscardovia omnicolens/Scardovia* sp. sp4-iso-1_F08 | 86 891 | 949.4 | NA | NBLEM | −24.31 | NA | 1.37 (−0.68 to 3.42) | 0.9309 |
| *Prevotella timonensis* | 209 252 | 2527 | NA | NBLEM | 1.52 (−0.34 to 3.38) | 0.3873 | 0.78 (−0.18 to 1.74) | 0.8048 |
| *Finegoldia magna* | 129 000 | 3832 | NA | NBLEM | −0.72 (−1.89 to 0.45) | 0.6083 | −0.27 (−0.83 to 0.29) | 0.9309 |
| *Sneathia amnii/Leptotrichia amnionii* | 350 117 | 1048 | NA | NBLEM | 1.63 (−4.51 to 7.76) | 0.8646 | −3.71 (−6.77 to −0.65) | 0.4624 |
| *Lactobacillus crispatus* (99%)/*Lactobacillus acidophilus* (99%) | 278 931 | 2748 | NA | NBLEM | −0.96 (−2.33 to 0.41) | 0.5340 | −0.63 (−1.36 to 0.11) | 0.7446 |
| Uncultured Lachnospiraceae bacterium (94%) | 219 344 | 652.7 | NA | NBLEM | −2.11 (−6.34 to 2.12) | 0.6906 | 0.35 (−1.93 to 2.63) | 0.9309 |
| *Gardnerella vaginalis* | 232 527 | 1803 | NA | NBLEM | 2.85 (0.48 to 5.22) | 0.1003 | −0.46 (−1.80 to 0.87) | 0.9309 |
| *Prevotella bivia* | 222 106 | 2483 | NA | NBLEM | −1.02 (−3.06 to 1.01) | 0.6906 | 0.68 (−0.40 to 1.75) | 0.9309 |
| *Lactobacillus iners* (99%)/*Ralstonia solanacearum* (99%) | 213 474 | 1976 | NA | NBLEM | −3.98 (−6.48 to −1.47) | 0.0321** | 0.16 (−1.05 to 1.37) | 0.9309 |
| *Lactobacillus jensenii* | 144 856 | 2989 | NA | NBLEM | −0.71 (−2.43 to 1.01) | 0.7569 | 0.16 (−0.73 to 1.04) | 0.9309 |

AIC, Akaike information criterion; NA, not applicable due to errors; NBLEM, negative binomial linear effect model; PLEM, Poisson linear effects model; ZINBLEM, zero-inflated negative binomial linear effect model.

**P*-value considered significant if <0.05.

***Significant *P*-values after adjustment for number of tests by false discovery rate (FDR) and confounders (maternal age, pre-pregnancy body mass index, ethnicity, parity and smoking history).

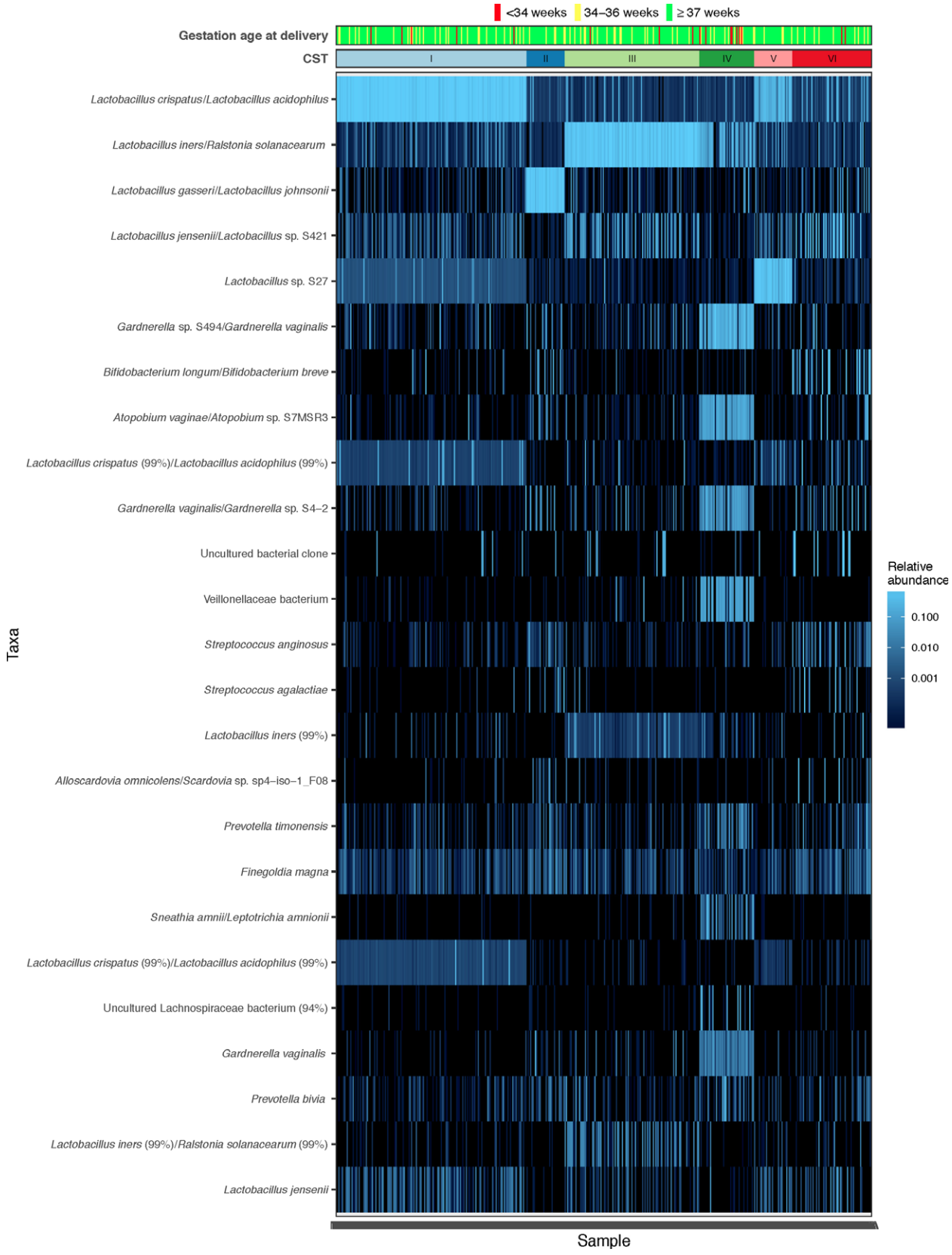

**Figure 1.** Heat map of the relative abundance of the 25 most abundant oligotypes, based on the MED algorithm, in the vaginal communities of 450 women sampled early in pregnancy. Clustering on the abundance profiles of samples using the partitioning around the medoids algorithm identified six community state types (CSTs). CSTs I, II, III and V were characterised by dominant *Lactobacillus* species: *Lactobacillus crispatus/Lactobacillus acidophilus*, *Lactobacillus gasseri/Lactobacillus johnsonii*, *Lactobacillus iners/Ralstonia solanacearum* and *Lactobacillus* sp. S27, respectively. One CST with greater diversity and presence of BV-associated bacteria (CST IV); and one CST with lower diversity and nondominance of *Lactobacillus* species (CST VI). Pregnancy outcomes are indicated by the bar at the top: early preterm birth (red) <34 weeks of gestation, late preterm birth (yellow) = 34–36 weeks of gestation term delivery (green) ≥37 weeks of gestation. The sample sizes within CST groups are as follows: $n$ = 160 (CST I), $n$ = 32 (CST II), $n$ = 113 (CST III), $n$ = 46 (CST IV), $n$ = 32 (CST V) and $n$ = 67 (CST VI), respectively.

increased risk of early but not late spontaneous preterm birth compared with a CST with low diversity and non-dominance of *Lactobacillus* species.

## Strengths

Previous studies on the association between vaginal microbiome and risk of preterm birth had smaller sample sizes ($n < 34$)[2,8–11] than our study ($n = 94$). Therefore, one of the strengths of our study is the higher power for detection of differences compared with previous studies.

## Limitations

The V1-V3[8,9] region of the bacterial 16S rRNA versus V4 is commonly used to assess *Lactobacillus* community composition but some previous studies[2,25] have used similar regions for vaginal microbiome analyses. The V4 variable region of the 16S rRNA gene provides strong discrimination between most bacterial species.[26] However, additional computational methods such as oligotyping[27] may be needed to precisely identify certain species, such as *L. crispatus*, as performed in our study. Selection of the V4 region of 16SrRNA may limit the comparability of our results to studies using other regions.

Microbial–host interactions, use of antibiotics, progesterone and probiotics were not considered in the association between vaginal microbiome and preterm birth,[9] which are limiting factors.

## Interpretation

Lack of association between Nugent score and vaginal microbiome composition in our study is surprising but has previously been reported.[28] Various factors can contribute to this discrepancy in findings, for example, different methods of Gram staining may result in different Nugent score inter-rater reliability[29] and low abundance of bacteria that are better represented by sequencing than Nugent score.[30]

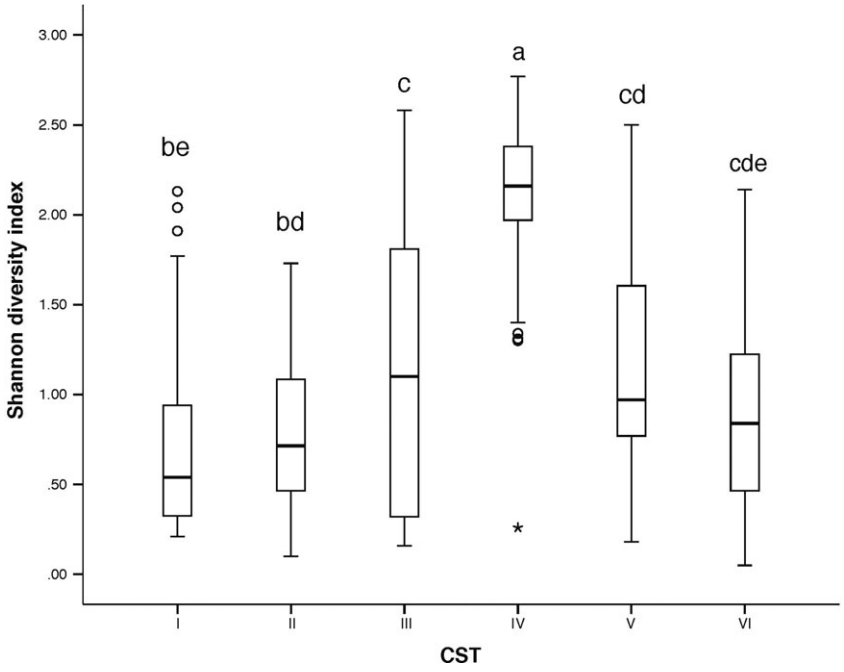


**Figure 2.** The microbial diversity (Shannon diversity index; SDI) between different community state types (CSTs). The SDI was significantly different between CSTs ($F = 44.26$, $P < 0.0001$). CST IV was significantly more diverse compared with all other CSTs ($P < 0.05$). Boxplots with different superscripts represent statistical significance ($P < 0.05$). The sample sizes within CST groups were as follows: $n$ = 160 (CST I), $n$ = 32 (CST II), $n$ = 113 (CST III), $n$ = 46 (CST IV), $n$ = 32 (CST V) and $n$ = 67 (CST VI), respectively.

 

Previous studies investigating the vaginal microbiome in term and preterm deliveries have reported inconsistent results.[2,8–11] DiGiulio et al.[2] reported a higher risk of preterm birth in patients within a CST poor in *Lactobacillus* but highly abundant in *Gardnerella* or *Ureaplasma*. Similarly, we observed an increased risk of early spontaneous preterm birth in patients from the CST IV (also highly abundant in *Gardnerella*) but only in comparison with the low-diverse CST (CST VI), which did not have dominance of *Lactobacillus*. The study population by DiGiulio et al. consisted of 63% white and 37% non-white women, which is relatively similar to our study. In their study,[2] vaginal swabs were collected weekly through pregnancy which is different from the early pregnancy collection of swabs in our study but the vaginal microbiome is known to be relatively stable during pregnancy[2,31] and mostly dominated by *Lactobacillus*.[7,32] However, patients with pregnancy complications, for example, pre-eclampsia[33] and fetal growth retardation,[34] are highly associated with preterm prelabour rupture of membranes, and therefore, nonspontaneous preterm births were excluded from our study but were included in DiGiulio's study. This discrepancy may account for different findings between studies.

Recent studies have suggested a protective role of *L. crispatus*[11,12,35] in the vaginal microbiota on preterm birth whereas others did not report any differences.[9] In our study, we observed that *L. crispatus* (99%)/*L. acidophilus* (99%) at early gestation may be associated with decreased risk of spontaneous preterm birth, which is similar to recent findings from two UK studies[11,35] (in predominantly white populations) and a study consisting of two cohorts of white and African American women.[12] Among other lactobacilli, *L. crispatus* has the largest genome,[36] and potential bacteriocin and adhesin genes.[37] Stafford et al.,[35] recently reported a positive association between *L. crispatus* and succinate levels. The latter may be protective against inflammation-associated preterm birth.[35] The observed inverse association between presence of *L. gasseri/L. johnsonii* and risk of early preterm birth in our study has been previously reported by Callahan et al. (in African American women).[12] However, the inverse association between relative abundance of *L. iners* (99%)/*R. solanacearum* (99%) and risk of early preterm birth is in contrast to previous studies[11,38] which reported a direct association between *L. iners* and risk of preterm birth. However, in our study, the relative abundance of *L. iners* (99%)/*R. solanacearum* (99%) operational taxonomic units is low compared with other significant operational taxonomic units (see Supplementary material, Figure S4C versus A, B and D). Also, additional statistical analyses suggest that the inverse association between presence of *L. gasseri/L. johnsonii* and *L. iners* (99%)/*R. solanacearum* (99%) is specific to women with early preterm (<34 weeks) and not preterm (<37 weeks) birth, whereas, *L. crispatus* (99%)/*L. acidophilus* (99%) has a protective effect on preterm birth when categorised as both (<34 and <37 weeks) (data not shown). *Bifidobacterium* are mainly abundant in the intestinal tract but are also detected in the vaginal tract.[39] Meta-analyses have not observed any association between consumption of *Bifidobacterium* probiotics during pregnancy and gestational age.[40] The observed protective association between *Bifidobacterium* and early preterm birth is interesting and requires further research.

Recently, a cross-sectional study from a longitudinal US cohort of predominantly (69%) African American women reported a lower level of SDI in preterm delivery and higher level of intercommunity diversity across pregnancy trimesters in women who delivered preterm versus term.[10] We observed significantly higher SDI in CST IV versus other CSTs and CST IV was associated with 4.2 times the risk of early spontaneous preterm birth compared with the low-diversity vaginal CST VI, which is different and not comparable to the stated study. However, in their study,[10] no distinct taxa at any pregnancy trimester was associated with preterm birth.[10] Perhaps lower power due to lower sample size of women at early pregnancy may have resulted in failure to observe differences between preterm and term deliveries at this pregnancy time-point,[10] which is an important phase of pregnancy in terms of the microbial community in association with risk of preterm birth.[2] Our results are also different from those of Hyman et al.. In their study, they reported no correlation between low abundance or absence of *Lactobacillus* throughout pregnancy and the risk of preterm birth, despite the lower bacterial diversity observed in preterm cases in the white subgroup.[8] Preterm birth included nonspontaneous preterm births at <37 weeks of gestation, which was different from the exclusive spontaneous preterm birth cases in our study. Our results also differ from a study that did not report any differences in bacterial taxa relative abundance and vaginal CSTs throughout pregnancy between women with spontaneous early preterm (<34 weeks) and term deliveries.[9] This study mainly included African Americans, which is different from the prominently white European population in our study.

Details of the interpretation of 'Distribution of vaginal community state types (CSTs) by ethnicity' can be found in the Supplementary material (Appendix S3).

## Conclusion

Our results suggest that *L. gasseri/L. johnsonii*, *L. crispatus/L. acidophilus*, *L .iners/R. solanacearum* and *B. longum/B. breve* may be associated with decreased risk of early but not late spontaneous preterm birth. Additionally, a vaginal CST with high diversity and presence of BV-associated bacteria (*G. vaginalis*, *A. vaginae* and *Veillonellaceae* bacterium)

may be associated with an increased risk of early but not late spontaneous preterm birth as compared with a CST with low diversity and nondominance of *Lactobacillus* species. Further studies exploring the association between the vaginal microbiome across pregnancy and risk of spontaneous preterm birth are recommended while considering the immunology of the host.

### Disclosure of interests

None declared. Completed disclosure of interests form available to view online as supporting information.

### Contribution to authorship

WDF and NT designed research; VY and AD performed the DNA extraction of the vaginal samples in LBB's laboratory; CA and NT performed statistical analyses of the data, bioinformatics analyses of the data was performed by AME; NT wrote the manuscript. AME, LBB and WDF provided constructive comments on the manuscripts. All authors read and approved the final manuscript.

### Details of ethics approval

The local institution as stated in the Methods section has approved human experimentation. Institutional Review Board Project #CHU-HSJ-2009-010 was approved on 10 June 2014.

### Funding

This study was conducted as part of the research programme of the Integrated Research Network in Perinatology of Quebec and Eastern Ontario (IRNPQEO). This work was supported by the Canadian Institutes of Health Research (CIHR) [CRI 88413]. WDF received the grant. The Postdoctoral Scholarship for NT was provided by QTNPR (Quebec Training Network in Perinatal Research) – CIHR (Canadian Institutes of Health Research).

### Acknowledgements

We would like to acknowledge all staff, researchers and patients of the 3D pregnancy cohort study for their great contribution to the study. We would like to thank Josée Poirier who was responsible for the 3D study biobank and preparing the aliquots of the biological samples for this study at the Research Centre of CHU Sainte-Justine. Acknowledgements are also due to the McGill University and Genome Quebec Innovation Centre for the MiSeq analyses of the extracted vaginal DNA samples.

## Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Figure S1.** The consort flow diagram for the study.

**Figure S2.** The microbial biodiversity (Shannon diversity index; SDI) between Nugent score categories.

**Figure S3.** The microbial biodiversity (Shannon diversity index; SDI) among overall Nugent scores.

**Figure S4.** Vaginal microbial oligotype relative abundance (logarithmic scale) composition between women who had early (<34 weeks of gestation) and late (34–36 weeks of gestation) preterm versus term (≥37 weeks of gestation) deliveries.

**Figure S5.** Vaginal microbial taxonomy relative abundance (logarithmic scale) composition of *Bifidobacterium* between women who had early (<34 weeks of gestation) and late (34–36 weeks of gestation) preterm versus term (≥37 weeks of gestation) deliveries.

**Table S1.** Primer sequences used for polymerase chain reaction (PCR) amplification of the V4 hypervariable regions of 16S rRNA genes.

**Table S2.** Vaginal microbial taxonomy differential relative abundance composition between pregnant women who had early (<34 weeks of gestation) and late (34–36 weeks of gestation) preterm versus term (≥37 weeks of gestation) deliveries.

**Table S3.** Association of vaginal community state types (CSTs) with early (<34 weeks) and late (34–36 weeks) spontaneous preterm birth versus term deliveries.

**Table S4.** Distribution of vaginal community state types (CSTs) by ethnicity.

**Appendix S1.** Supplementary methods.

**Appendix S2.** Supplementary results.

**Appendix S3.** Supplementary discussion. ■

## References

1 Riduan JM, Hillier SL, Utomo B, Wiknjosastro G, Linnan M, Kandun N. Bacterial vaginosis and prematurity in Indonesia: association in early and late pregnancy. *Am J Obstet Gynecol* 1993;169:175–8.

2 DiGiulio DB, Callahan BJ, McMurdie PJ, Costello EK, Lyell DJ, Robaczewska A, et al. Temporal and spatial variation of the human microbiota during pregnancy. *Proc Natl Acad Sci USA* 2015;112:11060–5.

3 Leitich H, Bodner-Adler B, Brunbauer M, Kaider A, Egarter C, Husslein P. Bacterial vaginosis as a risk factor for preterm delivery: a meta-analysis. *Am J Obstet Gynecol* 2003;189:139–47.

4 Nugent RP, Krohn MA, Hillier SL. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of Gram stain interpretation. *J Clin Microbiol* 1991;29:297–301.

5 Hilbert DW, Smith WL, Chadwick SG, Toner G, Mordechai E, Adelson ME, et al. Development and validation of a highly accurate quantitative real-time PCR assay for diagnosis of bacterial vaginosis. *J Clin Microbiol* 2016;54:1017–24.

6 Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, et al. Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci USA* 2011;15(108 Suppl 1):4680–7.

 

7 Aagaard K, Riehle K, Ma J, Segata N, Mistretta TA, Coarfa C, et al. A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. *PLoS One* 2012;7:e36466.
8 Hyman RW, Fukushima M, Jiang H, Fung E, Rand L, Johnson B, et al. Diversity of the vaginal microbiome correlates with preterm birth. *Reprod Sci* 2014;21:32–40.
9 Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Bieda J, et al. The vaginal microbiota of pregnant women who subsequently have spontaneous preterm labor and delivery and those with a normal delivery at term. *Microbiome* 2014;2:18.
10 Stout MJ, Zhou Y, Wylie KM, Tarr PI, Macones GA, Tuuli MG. Early pregnancy vaginal microbiome trends and preterm birth. *Am J Obstet Gynecol* 2017;217:356.e1–18.
11 Kindinger LM, Bennett PR, Lee YS, Marchesi JR, Smith A, Cacciatore S, et al. The interaction between vaginal microbiota, cervical length, and vaginal progesterone treatment for preterm birth risk. *Microbiome* 2017;5:6.
12 Callahan BJ, DiGiulio DB, Goltsman DSA, Sun CL, Costello EK, Jeganathan P, et al. Replication and refinement of a vaginal microbial signature of preterm birth in two racially distinct cohorts of US women. *Proc Natl Acad Sci USA* 2017;114:9966–71.
13 Gillespie SL, Christian LM, Neal JL. A proposed bio-panel to predict risk for spontaneous preterm birth among African American women. *Med Hypotheses* 2015;85:558–64.
14 Keelan JA, Payne MS. Vaginal microbiota during pregnancy: pathways of risk of preterm delivery in the absence of intrauterine infection? *Proc Natl Acad Sci USA* 2015;112:E6414.
15 Fraser WD, Shapiro GD, Audibert F, Dubois L, Pasquier JC, Julien P, et al. 3D cohort study: the Integrated Research Network in Perinatology of Quebec and Eastern Ontario. *Paediatr Perinat Epidemiol* 2016;30:623–32.
16 Donders GG, Van Calsteren K, Bellen G, Reybrouck R, Van den Bosch T, Riphagen I, et al. Predictive value for preterm birth of abnormal vaginal flora, bacterial vaginosis and aerobic vaginitis during the first trimester of pregnancy. *BJOG* 2009;116:1315–24.
17 Kindinger LM, MacIntyre DA, Lee YS, Marchesi JR, Smith A, McDonald JA, et al. Relationship between vaginal microbial dysbiosis, inflammation, and pregnancy outcomes in cervical cerclage. *Sci Transl Med* 2016;8:350ra102.
18 Robinson HP. Sonar measurement of fetal crown–rump length as means of assessing maturity in first trimester of pregnancy. *Br Med J* 1973;4:28–31.
19 Nelson DB, Bellamy S, Gray TS, Nachamkin I. Self-collected versus provider-collected vaginal swabs for the diagnosis of bacterial vaginosis: an assessment of validity and reliability. *J Clin Epidemiol* 2003;56:862–6.
20 Kashyap B, Singh R, Bhalla P, Arora R, Aggarwal A. Reliability of self-collected versus provider-collected vaginal swabs for the diagnosis of bacterial vaginosis. *Int J STD AIDS* 2008;19:510–3.
21 Eren AM, Vineis JH, Morrison HG, Sogin ML. A filtering method to generate high quality short reads using illumina paired-end technology. *PLoS One* 2013;8:e66643.
22 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 1995;57:289–300.
23 Aragon TJ. epitools: epidemiology tools. R package version 0.5-10. 2017 [https://CRAN.R-project.org/package=epitools]. Accessed 26 October 2017.
24 R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2016 [https://www.R-project.org/]. Accessed 12 December 2016.
25 Dareng EO, Ma B, Famooto AO, Adebamowo SN, Offiong RA, Olaniyan O, et al. Prevalent high-risk HPV infection and vaginal microbiota in Nigerian women. *Epidemiol Infect* 2016;144:123–37.
26 Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA* 2011;1:4516–22.
27 Eren AM, Maignien L, Sul WJ, Murphy LG, Grim SL, Morrison HG, et al. Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol Evol* 2013;4:1111–927.
28 Wessels JM, Lajoie J, Vitali D, Omollo K, Kimani J, Oyugi J, et al. Association of high-risk sexual behaviour with diversity of the vaginal microbiota and abundance of *Lactobacillus*. *PLoS One* 2017;12:e0187612.
29 Libman MD, Kramer M, Platt R. Comparison of Gram and Kopeloff stains in the diagnosis of bacterial vaginosis in pregnancy. *Diagn Microbiol Infect Dis* 2006;54:197–201.
30 Xiao B, Niu X, Han N, Wang B, Du P, Na R, et al. Predictive value of the composition of the vaginal microbiota in bacterial vaginosis, a dynamic study to identify recurrence-related flora. *Sci Rep* 2016;6:26674.
31 Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Nikita L, et al. Correction: The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. *Microbiome* 2014;2:10.
32 MacIntyre DA, Chandiramani M, Lee YS, Kindinger L, Smith A, Angelopoulos N, et al. The vaginal microbiome during pregnancy and the postpartum period in a European population. *Sci Rep* 2015;5:8988.
33 Auger N, Le TU, Park AL, Luo ZC. Association between maternal comorbidity and preterm birth by severity and clinical subtype: retrospective cohort study. *BMC Pregnancy Childbirth* 2011;11:67.
34 Kurzel RB, Lampley EC. Preterm premature rupture of the membranes is associated with asymmetric intrauterine growth retardation and smaller placentas. *Prim Care Update Ob Gyns* 1998;5:181.
35 Stafford GP, Parker JL, Amabebe E, Kistler J, Reynolds S, Stern V, et al. Spontaneous preterm birth is associated with differential expression of vaginal metabolites by lactobacilli-dominated microflora. *Front Physiol* 2017;8:615.
36 Witkin SS, Linhares IM. Why do lactobacilli dominate the human vaginal microbiota? *BJOG* 2017;124:606–11.
37 Ojala T, Kankainen M, Castro J, Cerca N, Edelman S, Westerlund-Wikström B, et al. Comparative genomics of *Lactobacillus crispatus* suggests novel mechanisms for the competitive exclusion of *Gardnerella vaginalis*. *BMC Genom* 2014;15:1070.
38 Petricevic L, Domig KJ, Nierscher FJ, Sandhofer MJ, Fidesser M, Krondorfer I, et al. Characterisation of the vaginal *Lactobacillus* microbiota associated with preterm delivery. *Sci Rep* 2014;4:5136.
39 Yang S, Reid G, Challis JR, Kim SO, Gloor GB, Bocking AD. Is there a role for probiotics in the prevention of preterm birth? *Front Immunol* 2015;6:62.
40 Dugoua JJ, Machado M, Zhu X, Chen X, Koren G, Einarson TR. Probiotic safety in pregnancy: a systematic review and meta-analysis of randomized controlled trials of *Lactobacillus*, *Bifidobacterium*, and *Saccharomyces* spp. *J Obstet Gynaecol Can* 2009;31:542–52.