

# Parity and gestational age are associated with vaginal microbiota composition in term and late term pregnancies

**Kaisa Kervinen**

University of Helsinki and Helsinki University Hospital <https://orcid.org/0000-0003-4195-556X>

**Tiina Holster**

University of Helsinki and Helsinki University Hospital

**Schahzad Saqib**

University of Helsinki

**Seppo Virtanen**

University of Helsinki and Helsinki University Hospital

**Vedran Stefanovic**

University of Helsinki and Helsinki University Hospital

**Leena Rahkonen**

University of Helsinki and Helsinki University Hospital

**Pekka Nieminen**

**Anne Salonen** (✉ [anne.salonen@helsinki.fi](mailto:anne.salonen@helsinki.fi))

University of Helsinki <https://orcid.org/0000-0002-6960-7447>

**Iikka Kalliala**

University of Helsinki <https://orcid.org/0000-0002-7664-0587>

---

## Article

**Keywords:** Microbiota, pregnancy, nulliparous, vagina, gestational age

**Posted Date:** November 1st, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1032686/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

## Abstract

Vaginal microbiota and its potential contribution to preterm birth has been under intense research in recent years. However, only few studies have studied vaginal microbiota in later stages of pregnancy or at the onset of labor. We analyzed vaginal swab samples collected between 37- and 42-weeks of gestation from 324 Finnish women before elective cesarean section, at the onset of spontaneous labor, and in pregnancies continuing beyond 41 weeks of gestation. Vaginal microbiota composition associated strongly with parity, i.e. previous deliveries, and advancing gestational age. Absence of previous deliveries was a strong predictor of *L. crispatus* dominated vaginal microbiota, and the relative abundance of *L. crispatus* was higher in late term pregnancies, especially among nulliparous women. The results underscore the importance of the vaginal microbiota for improving the currently limited understanding on how the duration of gestation and timing of birth is regulated, with potentially vast clinical utilities.

## Introduction

The vaginal microbial environment is affected and modified by multiple factors<sup>1–5</sup>. During pregnancy vaginal microbiota is often dominated by the genus *Lactobacillus*<sup>6</sup> and more stable and less diverse compared to the non-pregnant state<sup>7,8</sup>. Many factors have been hypothesized to explain this phenomenon, including the lack of cyclic hormonal fluctuations or menstruation, changes in cervicovaginal secretions, and decreased sexual activity<sup>7</sup>. Rising estrogen levels during pregnancy result in increasing vaginal glycogen accumulation, favoring the proliferation of lactobacilli<sup>9</sup> which protect the upper genital tract from ascending pathogens.

While the drastic changes in the vaginal microbiota on transition from pregnancy to post-partum state are well described<sup>10</sup>, only couple of studies have addressed the potential lasting effects of pregnancy and childbirth on woman's vaginal microbiota composition. An US-based study showed that number of previous pregnancies was associated with vaginal microbiota composition in the first trimester of pregnancy<sup>11</sup> while a recent large study in healthy Chinese non-pregnant, reproductive age women reported that pregnancy history is essentially the strongest influencer of the vaginal microbiota, surpassing the effect of e.g. menstrual cycle<sup>5</sup>. In both studies, dominance of *Lactobacillus crispatus* was found to be highest in women without previous conception.

Average human gestation lasts 40 weeks (280 days). Term pregnancy can be defined as early term (37 0/7 through 38 6/7 weeks of gestation (GW), full term (GW 39 0/7 through 40 6/7), late term (GW 41 0/7 through 41 6/7), and postterm (GW 42 0/7 and beyond)<sup>12</sup>. The vaginal microbiota differs between women giving birth prematurely or at term<sup>13,14</sup>. Studies on predominantly Caucasian populations have reported that increased diversity, dominance by *Lactobacillus iners* or overall depletion of *Lactobacillus* spp. increase the risk for preterm birth (PTB)<sup>15–18</sup>, whereas dominance of *Lactobacillus crispatus* in the vaginal microbiota would act as protection against PTB<sup>14,16,19</sup>. Vaginal microbiota has been suggested to change with increasing gestational age with the relative abundance of *Lactobacillus* spp. possibly

declining towards child birth<sup>20,21</sup>, and in late gestation to have similarities to the non-pregnant microbiota<sup>6</sup>.

The mechanisms behind the onset of spontaneous labor, as well as factors determining the duration of gestation remain largely unknown<sup>22</sup>. Late term and postterm pregnancy, i.e. gestation lasting more than 41 weeks is associated with maternal and fetal risks, higher rates of operative delivery, and increased perinatal mortality<sup>23,24</sup>. In Western countries, induction of labor is performed in almost 50% of all pregnancies that proceed beyond 41 weeks<sup>25</sup>, of which roughly 20% end up in cesarean section (CS)<sup>26</sup> with the risk being higher among nulliparous women, i.e. women with no previous deliveries<sup>27,28</sup>. However, studies investigating microbial composition with increasing duration of pregnancy are scarce<sup>21</sup>, and to our knowledge do not exist for Caucasian women with GW > 41 0/7.

Our objective was to study whether vaginal microbiota differs according to previous pregnancy outcomes and whether gestational age or other host-related factors influence the composition of vaginal microbiota at or near delivery.

## Results

### Cohort characteristics

A total of 324 pregnant women were included in the study, GW of which 155 (47.8%) gestational age was under 41.0 weeks and 169 (52.2%) equal to or exceed 41.0 weeks. Of the samples 60 (18.5%) were taken at elective CS (GW between 38.7–40.4), 96 at delivery (29.6%, GW between 37.1–41.0) and 168 (51.8%) at postterm visit at maternal outpatient clinic (GW between 41.0–42.0). The demographic, lifestyle and clinical characteristics of the study population are shown in Table 1. Of all studied women 58.0% (n=188) were nulliparous and of these 139 (73.9%) had no previous pregnancies and 49 (26.1%) had at least one miscarriage or medical abortion in their pregnancy history. The median gestational age at sampling was 41.0 weeks (IQR 39.7–41.7), the mean age 32.1 years (SD 4.8), and the median BMI 23.1 kg/m<sup>2</sup> (IQR 21.2–26.4).

### Vaginal microbiota in late pregnancy in Caucasian women

From this point forth, “presence/detection” of a bacterium is defined as >5% and “dominance” is defined as >50% of relative abundance in the sample. Lactobacilli were detected in 309/324 (95.4%) women, of which 256/324 (79.0%) had *Lactobacillus*-dominated vaginal microbiota (Fig. 1). *L. crispatus* was the most abundant and prevalent species overall, being present in 170/324 (52.5%) of the samples and the dominant bacterium of the vaginal microbiota of 142/324 (43.8%) of women. *L. iners* was detected in 119/324 (36.7%) and was the dominant species in 90/324 (27.8%) women. Other *Lactobacillus* species of note were *L. jensenii* (detected: 48/324, 14.8%; dominant: 12/324, 3.7%) and *L. gasseri* (detected: 31/324; 9.6%, dominant: 12/324, 3.7%). *Gardnerella* spp. were the most prevalent non-*Lactobacillus* bacteria and were present in 81/324 (25%) and dominant in 30/324 (9.3%) women. Two species of

*Gardnerella* were seen, *G. vaginalis* (detected: 78/324, 24.1%; dominant: 29/324, 9.0%) and *G. leopoldii* (detected: 18/324, 5.6%; dominant: 1/324, 0.3%). Altogether 55 different species-level taxa were detected, representing 24 genera and 7 higher taxonomic ranks (Supplementary Data 1).

## Vaginal microbiota and parity

The gross vaginal microbiota composition associated strongly with parity (permutational ANOVA F=21.30; R<sup>2</sup>=0.062; p<1.00e-05, Supplementary Table S5). The bacterial profiles of samples stratified by parity are shown in Fig. 1. *L. crispatus* was more frequent in nulliparas (detected: 123/188, 65.4%; dominant: 111/188, 59.0%) than in multiparas (detected: 47/136, 34.6%; dominant: 31/136, 22.8%) (Fig. 1, Supplementary Table S3-S4). The mean relative abundance of *L. crispatus* was 2.5 times more abundant in samples from nulliparous women compared to multipara (58.2% vs. 23.3%, q<0.001, Fig. 2, Supplementary Table S6). On the contrary, *L. iners*, *L. gasseri*, *G. vaginalis*, and *Famyhessea vaginæ* (previously *Atopobium vaginæ*) were more commonly seen in multiparas (33.5% in multiparas vs. 21.8% in nulliparas, q=0.03; 8.48% vs. 1.38%, q<0.001; 14.3% vs. 7.55%, q>0.05; 1.89% vs. 0.64%, q>0.05, respectively).

To study the effect of previous pregnancies vs. deliveries, the nulliparous women were further stratified into primigravidas (no previous pregnancies) and nulliparous multigravidas (presence of previous pregnancies that ended in spontaneous or induced abortion). No significant differences in relative abundances of any bacteria were observed between these two groups (Fig. 2, Supplementary Fig. 1, Supplementary Table S7-S8). *L. crispatus* was the most abundant bacterium overall but was detected more often in nulliparous multigravidas (mean relative abundance nulliparous multigravidas: 66.5% vs. primigravidas: 55.2%). The distribution of *L. crispatus* in both stratified subgroups of nulliparas compared to multiparas was consistently significant (q<0.001) as was that of *L. gasseri* (mean relative abundance primigravidas: 1.03%, q<0.001; nulliparous multigravidas: 2.37%, q=0.01). The mean relative abundance of *L. iners* in nulliparous multigravidas (16.0%) was lower than in the primigravidas' samples (23.9%, q=0.19) and significantly lower than among multiparas (q=0.01). *F. vaginæ* was not observed in nulliparous multigravidas' samples (q=0.01).

In women with one prior delivery compared to nulliparas, the differences in the relative abundances of *L. crispatus* and *L. gasseri* were only observed after vaginal delivery or birth by emergency cesarean section (q<0.001), whereas no significant differences in their abundances were seen between nulliparas and women with one previous delivery by elective cesarean section (Supplementary Fig. 2, Supplementary Table S9). However, the sample size of women with one previous delivery by elective cesarean section was small (n=8).

The number of previous deliveries correlated negatively with abundance of *L. crispatus*: the mean relative abundance decreased from 58.1% in nulliparas to 25.7% in women with one prior delivery and to 15.4% with two or more deliveries (q<0.001, Supplementary Fig. 3, Supplementary Table S10). Respectively, the relative abundances of *L. gasseri* and *L. iners* increased with rising number of previous deliveries (q<0.001 and q=0.046, respectively, Supplementary Table S10). *G. vaginalis* was less abundant (q=0.04)

the longer the interval between previous delivery and current pregnancy, but no differences in *Lactobacillus* abundances were observed in relation to time since last delivery (Supplementary Table S11).

## Vaginal microbiota and increasing gestational weeks

Gestational age at the time of sampling associated significantly with microbiota variation ( $F=4.44$ ,  $R^2=0.014$ ,  $p=0.007$ , Supplementary Table S5). *L. crispatus* was the most abundant bacterium in both term (GW <41.0) and late term (GW  $\geq$ 41.0) samples (mean relative abundance 37.5% and 49.0%, respectively,  $q=0.33$ ), but a significant increase in the relative abundance of *L. crispatus* was seen with advancing gestational age ( $q=0.04$ , Fig. 3, Supplementary Table S12).

When stratified according to parity, the rise in the abundance of *L. crispatus* with increasing gestational age was only evident in nulliparas ( $q=0.04$ ), not in multiparas (Fig. 3, Supplementary Table S13), and more precisely among primigravidas ( $q=0.02$ , Supplementary Table S14). Among primigravidas the relative abundance of *L. gasseri* increased ( $q<0.001$ ) while the abundance of *L. iners* and *L. jensenii* declined along gestational age ( $q=0.03$  and  $q=0.02$ , respectively), whereas for nulliparous multigravidas the amount of *L. gasseri* significantly decreased ( $q<0.001$ ). No differences in bacterial relative abundances were observed among multiparas with rising gestational age (Supplementary Table S13).

The parity related differences in the vaginal microbiota were more distinct in late term samples compared to samples taken at term with significant differences in *L. crispatus* ( $q<0.001$ ), *L. gasseri* ( $q<0.001$ ), *L. iners* ( $q=0.04$ ), *G. vaginalis* ( $q=0.04$ ) and *F. vaginae* ( $q=0.03$ ) relative abundances between nulliparas and multiparas (Fig. 4, Supplementary Table S15-S16).

## Relationships between microbiological and demographic characteristics

The relationships of demographic variables with microbiota variation and single bacterial taxa are shown in Supplementary File 2 – Tables S17-S25. Education level correlated significantly with the overall microbiota variation ( $F=4.31$ ,  $R^2=0.014$ ,  $p=0.007$ , Supplementary Table S5) and higher education was associated with lower relative abundance of *L. iners* ( $q<0.001$ ), *G. vaginalis* ( $q=0.005$ ) and *G. leopolidii* ( $q<0.001$ ). Smoking was associated with higher abundance of *L. iners* ( $q=0.004$ ) as was higher amount of lifetime sex partners ( $q=0.001$ ), whereas higher BMI associated with higher abundance of *G. vaginalis* ( $q=0.02$ ). Smoking was also related to lower abundance of *L. gasseri* ( $q<0.001$ ). History of fertility treatments either related to current pregnancy or in the past was associated with lower prevalence of *F. vaginae* ( $q<0.001$ ). Intercourse <48 hours prior to sampling associated with lower abundance of *L. gasseri* ( $q=0.01$ ). Use of probiotics associated with lower levels of *G. leopolidii* ( $q=0.003$ ). The groups in which the samples were collected (elective cesarean section, during delivery, first postterm antenatal visit) did not correlate with gross microbiota variation ( $F=1.23$ ,  $R^2=0.008$ ,  $p=0.27$ , Supplementary Table S5) and no differences in abundances of bacteria between groups was seen. Maternal age, previous preterm births, gestational diabetes in current pregnancy, reported gynecological infections, use of antibiotics in

the recent three months, postpartum infections, or contractions at the time of sample did not associate significantly with the overall microbiota variation (Supplementary Table S5) or abundance of any individual bacterium.

## Discussion

The characteristics and individual variations of the bacterial microbiota in the human reproductive tract are well described but little is known about the influencing factors. Here, we showed in a relatively large cohort of Caucasian pregnant women that both the reproductive history as well as gestational age strongly affect the abundance and prevalence of the dominant vaginal bacteria, many of which have well-established associations to gynecological and reproductive health. We observed that nulliparity associated strongly with *L. crispatus*-dominated vaginal microbiota in term and late term pregnancies. Prior pregnancies ending in spontaneous or induced abortion did not alter this association. We also demonstrated that the mode of delivery might have an impact on the vaginal microbiota as among women with one previous delivery, in those with a history of elective CS, the vaginal microbiota corresponded to that of a nulliparous woman, while after one either vaginal delivery or delivery by emergency CS the microbiota diverged. On the other hand, increasing number of prior deliveries associated with decreasing prevalence of *L. crispatus*. Overall, the vaginal microbiota in the late third trimester varied according to the duration of gestation, and especially among women with no previous deliveries *L. crispatus* was more prevalent the longer the duration of the pregnancy was.

## Interpretation of the results in light of previous evidence

We found that the vaginal microbiota differed profoundly between nulliparous and multiparous women, indicating that the reproductive history is reflected in the vaginal microbiota at or close to delivery. Our results confirm and extend similar findings reported earlier for the first trimester of pregnancy (between 8- and 12-weeks of gestation)<sup>11</sup>, and for non-pregnant reproductive aged women<sup>5</sup>. The depletion of vaginal lactobacilli postpartum compared to pregnancy is well established and coincides with the parturition-induced drop in estrogen levels<sup>19,29</sup>. However, the duration of *Lactobacillus*-deficient postpartum microbiota signature has not been thoroughly studied, though there are indications that it remains up to one year postpartum<sup>30</sup>. Recent research on US-based cohort suggests that such signature is also present in women who delivered by an unlabored cesarean surgery<sup>31</sup>, while another recent study identified birth mode-dependent differences on the vaginal microbiota of Chinese women sampled at comparable time 6 weeks postpartum<sup>32</sup>.

A successful pregnancy requires tightly coordinated and balanced interplay between host innate immune defenses, mucosal immune responses, and the resident microbiota<sup>33,34</sup>. Although we studied only vaginal microbiota, the inflammatory and/or adaptive and innate immunity profiles are also anticipated to differ based on pregnancy history. For instance, immunosuppressive regulatory T cells (FOXP3+ CD4 T lymphocytes), which recognize paternal antigens and are essential for maternal fetal tolerance,

accumulate during the first pregnancy and persist to some level in the maternal circulation after delivery<sup>35,36</sup>. Later in the subsequent pregnancies these cell populations re-expand in a quicker manner compared with the initial pregnancy<sup>35</sup>. A pregnancy-induced memory cell response and pregnancy alloimmunization has also been proposed to be behind higher transplantation graft rejection in women with prior pregnancies<sup>37</sup>.

As a further support for immunological rather than direct and local microbiological effects, parity has been shown to affect not only the vaginal microbiota but also gut microbiota during subsequent gestation in a pig model<sup>38</sup>. This lends further support for the immunological effects as a mechanism for persisting microbiota changes as a result of previous delivery. Hence, our parity-related microbiota findings could reflect the state of the local inflammatory processes and potentially suggest some form of an immunological memory from prior labor. This is also supported by our suggestive finding that only previous vaginal delivery or emergency CS, both involving the physiological process of labor, were associated with changes in vaginal microbiota composition. We observed no differences in the relative abundances of *L. crispatus* or *L. gasseri* after elective CS, albeit the sample size in this subgroup was small. A recent study composed of healthy Chinese non-pregnant, reproductive age women<sup>5</sup> showed that the vaginal microbiota is significantly associated with past reproductive events. Notably, questionnaire-based information on number of pregnancies, pregnancy history, mode of recent delivery, and number of deliveries were shown to be significantly associated with the microbiota, especially *L. crispatus*, corroborating our findings. Although intriguing, the hypothesis on the role of immunological memory warrants further research e.g. on the immunologic pathways activated during labor, and their possible differences between nulliparous and multiparous women. These studies would also help to elucidate whether the lower relative abundance of *L. crispatus* is a primary or secondary effect of previous delivery i.e., a persisting postpartum bacterial signature or reflecting an immunological imprint from the previous parturition process.

The ultimate physiological stimuli leading to the onset of labor remain unclear despite intensive research. Although *Lactobacillus*-depleted vaginal microbiota has been associated with increased risk of PTB<sup>39</sup>, studies on microbiota composition at term and especially at late term and prolonged gestation are scarce. The duration of parturition is usually shorter in women with prior deliveries compared to nulliparous<sup>40</sup>, and nulliparity is a known risk factor for prolonged gestation<sup>41</sup>. We also observed that *L. crispatus* dominance increased along the progression of gestational weeks at late pregnancy, especially among nulliparas. This could be explained by higher estrogen levels at late pregnancy as estrogen concentrations increase with gestational age<sup>42</sup>, and due to the higher estrogen levels during the whole pregnancy among nulliparas than multiparas<sup>43</sup>. Our results support the data of Romero et al.<sup>44</sup> who showed in their longitudinal study that the relative abundance of *Lactobacillus* spp., including *L. crispatus*, increased as a function of gestational age. Dissimilar observations have, however, been presented as Avershina et al.<sup>21</sup> first reported an increasing diversity of the vaginal microbiota at the onset of labor compared to samples taken at 36 weeks of gestation. Later, Rasmussen et al.<sup>20</sup> observed that there was a gradual decline of *Lactobacillus* spp. from week 24 of pregnancy until birth while only genera

*Enterococcus* and *Granulicatella* were associated with gestational age at birth. Their sample size (n=57) was small compared to ours and the last samples taken later during birth after rupture of membranes so the results might not be thoroughly comparable. Furthermore, the proportion of nulliparous women, in whom the association between higher gestational age and *L. crispatus* dominance was more pronounced in our study, was lower in their study (40% vs 58%).

Ripening of the cervix in normal parturition is characterized by inflammatory changes, including the activation of leucocytes and increasing levels of pro-inflammatory cytokines which remodel the extracellular matrix of the cervix<sup>45–47</sup>. An increased ratio of L- to D-lactic acid isomers may alter the cervical tissue integrity by activating matrix metalloprotein inducers<sup>29,48</sup>. Low concentrations of D-lactate versus L-lactate have been seen in association with *L. iners* and *G. vaginalis*<sup>48</sup>, whereas concentrations of D-lactate are higher in *L. crispatus*-dominated vaginal microbiota, reflecting the known metabolic characteristics of vaginal *Lactobacillus* species<sup>49</sup>. In *in vitro* studies an inflammatory response evoked by *L. crispatus* has been shown to be weaker than by *L. iners* or anaerobic bacteria<sup>50</sup>. Hence, our findings on the high abundance of *L. crispatus* in late and postterm pregnancies may reflect decreased inflammatory bacterial signals. While these findings from a cross-sectional study cannot address causality, it is tempting to hypothesize that the vaginal microbiota and its interactions with the host immune system could play a role in the maintenance of gestation and initiation of spontaneous term or late term labor.

Identifying women susceptible to prolongation of pregnancy and its associated complications would be of great clinical benefit. Prolonged pregnancy increases the risk of stillbirth, neonatal morbidity and mortality, and the risk of cesarean delivery associated to the induction of labor is higher than in inductions in earlier pregnancy weeks<sup>51</sup>. Nulliparity is one of the known risk factors for prolonged pregnancy<sup>52</sup>, and nulliparas form the majority of women undergoing induction of labor due to prolonged pregnancy<sup>28</sup>. Previously *Lactobacillus* abundances have been shown to decline gradually toward birth<sup>20</sup>. Prevalence of *L. crispatus*-dominated vaginal microbiota, however, was higher after the due date among nulliparas in our study. Although this finding alone is not sufficient to predict the duration of gestation, more detailed research in inflammatory and immunological pathway activation in addition to the microbiota composition could well help to identify women at risk of prolonged gestation. Also, as different clinical scoring systems to predict the success of labor induction have been proposed<sup>53,54</sup>, we may speculate that the vaginal microbiota could potentially be included in the evaluation to decrease the rates of unsuccessful labor inductions which could decrease maternal and neonatal complications and affect maternal birth experience.

The strengths of our study include the prospective setting and relatively high number of study subjects and the use of modern sequencing techniques for analyzing the microbiota. The study population was homogenic with comprehensive outcome and background information from medical records and the specific questionnaire designed for the study. To our knowledge, this study is the first to characterize the vaginal microbiota at late pregnancy stages in a large cohort of women with covering obstetric records. The limitations include sampling at a single timepoint for each participant. A longitudinal study with

samples from early pregnancy until delivery would allow us to see the potential intraindividual changes in the microbiota with advancing gestation. The homogeneity of our study population, despite it being also an asset, does not let us interpret whether the results can be generalized to women with different ethnic and biogeographical backgrounds knowing that ethnicity-related differences in the vaginal microbiota exist also during pregnancy<sup>29,55</sup>.

In conclusion, previous pregnancy history and mode of previous delivery have a strong association with the composition of vaginal microbiota close to parturition, while among nulliparous women, gestational age associates with the microbiota. The findings are novel and add to the few currently known factors associated with the duration of gestation (e.g. the relationship of maternal age, nulliparity and obesity to postterm pregnancy)<sup>41</sup>. Whether the microbiota changes just reflect or actively contribute to the underlying immunological processes and mechanisms remain to be studied. As the vaginal microbiota is, at least conceptually, modifiable, our findings highlight the importance of future studies to understand the nature of implicated host-microbiota interactions and to investigate its potential for diagnostic and therapeutic approaches.

## Methods

The study was carried out at the Department of Obstetrics and Gynecology, Helsinki University Hospital, Helsinki, Finland between May 2017 and December 2018. The study was approved by the ethical committee of the Hospital District of Helsinki and Uusimaa (HUS/907/2017) and performed in accordance with the principles of the Helsinki Declaration. All participants signed an informed consent and participation was voluntary.

A flow chart for selection of study population is shown in Supplementary Fig. 4. Exclusion criteria included multifetal gestation, maternal type 1 and type 2 diabetes, placenta-related pregnancy complications (e.g. pre-eclampsia and/or intrauterine growth restriction), and maternal bloodborne infectious diseases (viral hepatitis, HIV). The final study population included 324 women.

The participants were asked to complete a background questionnaire including questions about gynecological history, sexual habits, previous infections, antibiotic and probiotic use, relationship status, and educational status. Maternal pre-pregnancy body mass index (BMI), pregnancy history, smoking habits (current and/or previous smoking), postpartum infections and gestational age at sampling were obtained from antenatal clinic charts. A pregnancy with gestational age at sampling of under 41 weeks of gestation was considered term and of 41 weeks and beyond late term. Gestational age was determined by the crown-rump-length measurement at the time of the first trimester ultrasound screening. A woman with prior deliveries was considered multiparous irrespective of the number of previous deliveries and a woman with no deliveries in history was considered nulliparous. Further, nulliparous was considered primigravida if she had had no previous pregnancies and nulliparous multigravida if she had spontaneous/induced abortion(s) in history. Degree of education was reported as five-class (comprehensive school, vocational upper secondary education, technical college, general upper

secondary education, tertiary degree from university or university of applied sciences). A tertiary degree was defined as higher education. The number of sex partners during lifetime was reported as dichotomous (more than three partners or less). Postpartum infections were defined as infections requiring antibiotic treatment within two weeks after delivery and comprised of endometritis, episiotomy infections, post-cesarean wound infections and urinary tract infections. Time since previous pregnancy was counted in months between the sample and end date of previous pregnancy, either the date of previous delivery or the date when previous abortion was diagnosed.

Mid vaginal wall samples were collected with sterile flocked swabs (FLOQSwabs, Copan spa®, Italy) by healthcare professionals at the time of planned elective cesarean section, on admittance to the delivery ward due to contractions (i.e. women at the first stage of labor), or at the time of appointment when pregnancy proceeded beyond due date. According to the current departmental management guidelines, all women with an uncomplicated pregnancy receive an appointment for an antenatal visit in the maternity outpatient clinic latest at 41 5/7 weeks of gestation. All women had intact fetal membranes at the time of sampling. Tips of the swabs were severed to 1.5 mL Eppendorf tubes which were frozen in -20°C immediately after sampling and further moved to -80°C within two weeks.

## **DNA extraction and sequencing of the 16S rRNA gene amplicons**

Bacterial DNA was extracted from the swabs using a bead beating method and subjected for quality control, index PCR and Illumina MiSeq sequencing of the V3-V4 region of the 16S rRNA gene using primers 341F 5'-CCTACGGNGGCWGCAG-3' and 785Rev 5'-GACTACHVGGGTATCTAATCC-3') as previously described<sup>4</sup>. Samples were sequenced in four different runs, the samples from different study groups being randomly assigned to the runs. Part of the samples were sequenced together with fungal amplicons; in that case the bacterial reads were first extracted from the fastq-files with cutadapt based on primer sequences<sup>56</sup>. No run ID based batch effects on the microbiota were observed in permutational ANOVA (F1.11, R2= 0.010, p=0.33).

## **Sequence preprocessing and analysis**

The paired-end sequencing data was pre-processed using the dada2 R package (1.16) and its accompanying workflow (tutorial 1.16). The amplicon sequence variants (ASVs) obtained from dada2 were annotated using the taxminer R package workflow<sup>57</sup>. This consists of BLAST based sequence alignments followed by a text-mining based strategy that attaches ecosystem specificity to each alignment (host & isolation source) to probabilistically select the most likely taxonomic annotations.

## **Statistical analysis**

Permutational multivariate analysis of variance (PERMANOVA; adonis function in the vegan package<sup>58</sup> based on the Bray-Curtis distance was used to identify factors that explain between-sample variation in the microbiota (beta-diversity, Supplementary Table S5). Comparison of categorical variables was

measured with two-tailed Chi square and Fisher's exact test when appropriate, and of continuous variables with two-tailed t-test when normally distributed and two-tailed Mann Whitney U-test for non-normally distributed variables. GroupTest and CovariateTest from the mare R package<sup>59</sup> were used for species-wise comparisons between subgroups within the cohort, and for studying associations between clinical variables and individual taxa, respectively. These functions sequentially apply appropriate transformations and select the optimal models and parameters for each taxon based on the data distributions and model performance/fit. Some of the models being used include lm and glm from the stats package<sup>60</sup>, glm.nb from the MASS package<sup>61</sup> and gls from the nlme package<sup>62</sup>. If no model is found, where the data meets the model assumptions, no *p*-value is reported. A minimum acceptable prevalence and relative abundance of a taxon was set to 5%. The statistical models of mare functions use sample read count as an offset and *p*-values are corrected for false discovery rate (FDR; Benjamini-Hochberg)<sup>63</sup>. FDR-corrected *p*-values are reported in the text as *q*-values, and a value <0.05 was considered statistically significant. Biologically relevant background variables such as probiotic use and BMI were tested as potential confounders using PERMANOVA analysis. All figures were created in Rstudio using ggplot, gghalves, and cowplot R packages.

## Declarations

### DATA AVAILABILITY

All relevant data supporting the key findings of this study are available within the article and its Supplementary Information files, or from the corresponding author upon reasonable request. The sequencing data generated in this study have been deposited to the European Nucleotide Archive (ENA) with the project accession number PRJEB47492 and will be released upon acceptance.

### CODE AVAILABILITY

The code scripts used for data processing, analysis, and visualization have been deposited to GitHub (<https://github.com/SchahzadSaqib/EMV>).

### ACKNOWLEDGEMENTS

We want to thank all patients who volunteered to take part in our study, the midwives and doctors at Helsinki University Central Hospital who helped us to collect the samples, and Tinja Kanerva for processing the samples for sequencing. This study was supported by grants from EU H2020 programme grant number 814102 Sweet Crosstalk ITN to AS, from the Academy of Finland, Clinical Researcher grant to IK, State Funded Grant to PN, from the Finnish Medical Foundation to KK and TH, from Orion Research Foundation to KK, from Doctoral School in Health Sciences, University of Helsinki and Helsinki University Hospital to KK and SV.

### AUTHOR CONTRIBUTIONS

KK, IK, AS and PN conceptualized and designed the study. Patient recruitment, enrollment, and sample collection: KK, IK, TH. Analysis and interpretation of data: KK, SS, TH, SV, AS, IK. Writing the first version of the manuscript: KK, TH. Writing, review, and revision of the manuscript: KK, SS, IK, TH, LR, VS, AS, PN. All authors read, edited, revised, and approved the final version of the manuscript. These authors contributed equally: Tiina Holster and Schahzad Saqib; Anne Salonen and Ilkka Kalliala.

## COMPETING INTERESTS

The authors declare no competing interests.

## MATERIALS & CORRESPONDENCE

Correspondence to Anne Salonen.

## DECLARATIONS

Ethics approval and consent to participate

The study was approved by the ethical committee of The Hospital District of Helsinki and Uusimaa and Helsinki region hospital district (HUS/907/2017) and performed in accordance with the principles of the Helsinki Declaration. All participants signed an informed consent.

## SUPPLEMENTARY INFORMATION

Supplementary information can be found in separate files, Supplementary information.pdf, Description of additional supplementary files.docx and Supplementary Data 1.xlsx.

## References

1. Gajer, P. et al. Temporal dynamics of the human vaginal microbiota. *Sci. Transl. Med.* **4**, 132ra52 (2012).
2. Ravel, J. et al. Vaginal microbiome of reproductive-age women. *Proc. Natl. Acad. Sci. U. S. A.* **108 Suppl**, 4680–7 (2011).
3. Fettweis, J. M. et al. Differences in vaginal microbiome in African American women versus women of European ancestry. *Microbiology* **160**, 2272–82 (2014).
4. Virtanen, S. et al. Vaginal Microbiota Composition Correlates Between Pap Smear Microscopy and Next Generation Sequencing and Associates to Socioeconomic Status. *Sci. Rep.* **9**, 7750 (2019).
5. Jie, Z. et al. Life History Recorded in the Vagino-cervical Microbiome Along with Multi-omics. *Genomics Proteomics Bioinforma.* (2021) doi:10.1016/j.gpb.2021.01.005.
6. Aagaard, K. et al. A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. *PLoS One* **7**, e36466 (2012).

7. Walther-António, M. R. S. et al. Pregnancy's Stronghold on the Vaginal Microbiome. *PLoS One* **9**, e98514 (2014).
8. Romero, R. et al. The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. *Microbiome* **2**, 4 (2014).
9. Mirmomsef, P. et al. Free glycogen in vaginal fluids is associated with Lactobacillus colonization and low vaginal pH. *PLoS One* **9**, e102467 (2014).
10. Nunn, K. L. et al. Changes in the Vaginal Microbiome during the Pregnancy to Postpartum Transition. *Reprod. Sci.* **28**, 1996–2005 (2021).
11. Nasioudis, D. et al. Influence of Pregnancy History on the Vaginal Microbiome of Pregnant Women in their First Trimester. *Sci. Rep.* **7**, 10201 (2017).
12. ACOG Committee Opinion No 579: Definition of term pregnancy. *Obstet. Gynecol.* **122**, 1139–1140 (2013).
13. Di Simone, N. et al. Recent Insights on the Maternal Microbiota: Impact on Pregnancy Outcomes. *Front. Immunol.* **11**, 528202 (2020).
14. Kindinger, L. M. et al. The interaction between vaginal microbiota, cervical length, and vaginal progesterone treatment for preterm birth risk. *Microbiome* **5**, 6 (2017).
15. Brown, R. G. et al. Vaginal dysbiosis increases risk of preterm fetal membrane rupture, neonatal sepsis and is exacerbated by erythromycin. *BMC Med.* **16**, 9 (2018).
16. Tabatabaei, N. et al. Vaginal microbiome in early pregnancy and subsequent risk of spontaneous preterm birth: a case-control study. *BJOG* **126**, 349–358 (2019).
17. Freitas, A. C., Bocking, A., Hill, J. E., Money, D. M. & VOGUE Research Group, the V. R. Increased richness and diversity of the vaginal microbiota and spontaneous preterm birth. *Microbiome* **6**, 117 (2018).
18. Callahan, B. J. 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. U. S. A.* **114**, 9966–9971 (2017).
19. Kervinen, K. et al. Vaginal microbiota in pregnancy: Role in induction of labor and seeding the neonate's microbiota? *J. Biosci.* **44**, 116 (2019).
20. Rasmussen, M. A. et al. Ecological succession in the vaginal microbiota during pregnancy and birth. *ISME J.* **14**, 2325–2335 (2020).
21. Avershina, E. et al. Diversity of vaginal microbiota increases by the time of labor onset. *Sci. Rep.* **7**, 17558 (2017).
22. Kamel, R. M. The onset of human parturition. *Arch. Gynecol. Obstet.* **281**, 975–82 (2010).
23. Olesen, A. W., Westergaard, J. G. & Olsen, J. Perinatal and maternal complications related to postterm delivery: a national register-based study, 1978-1993. *Am. J. Obstet. Gynecol.* **189**, 222–7 (2003).
24. Caughey, A. B., Stotland, N. E., Washington, A. E. & Escobar, G. J. Maternal and obstetric complications of pregnancy are associated with increasing gestational age at term. *Am. J. Obstet. Gynecol.* **196**, 155.e1–6 (2007).

25. Wolff, S. L. et al. Has perinatal outcome improved after introduction of a guideline in favour of routine induction and increased surveillance prior to 42 weeks of gestation?: A cross-sectional population-based registry study. *Sex. Reprod. Healthc.* **10**, 19–24 (2016).
26. Alexander, J. M., McIntire, D. D. & Leveno, K. J. Prolonged pregnancy: induction of labor and cesarean births. *Obstet. Gynecol.* **97**, 911–5 (2001).
27. Vahrtian, A., Zhang, J., Troendle, J. F., Sciscione, A. C. & Hoffman, M. K. Labor Progression and Risk of Cesarean Delivery in Electively Induced Nulliparas. *Obstet. Gynecol.* **105**, 698–704 (2005).
28. Kruit, H. et al. Management of prolonged pregnancy by induction with a Foley catheter. *Acta Obstet. Gynecol. Scand.* **94**, 608–614 (2015).
29. MacIntyre, D. A. et al. The vaginal microbiome during pregnancy and the postpartum period in a European population. *Sci. Rep.* **5**, 8988 (2015).
30. Di Giulio, D. B. et al. Temporal and spatial variation of the human microbiota during pregnancy. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11060–5 (2015).
31. Pace, R. M. et al. Complex species and strain ecology of the vaginal microbiome from pregnancy to postpartum and association with preterm birth. *Med* **2**, 1027-1049.e7 (2021).
32. Zhang, X. et al. Variation of the vaginal microbiome during and after pregnancy in Chinese women. *medRxiv* 2020.07.07.20148536 (2020) doi:10.1101/2020.07.07.20148536.
33. Mei, C., Yang, W., Wei, X., Wu, K. & Huang, D. The Unique Microbiome and Innate Immunity During Pregnancy. *Front. Immunol.* **10**, 2886 (2019).
34. Doerflinger, S. Y., Throop, A. L. & Herbst-Kralovetz, M. M. Bacteria in the vaginal microbiome alter the innate immune response and barrier properties of the human vaginal epithelia in a species-specific manner. *J. Infect. Dis.* **209**, 1989–99 (2014).
35. Rowe, J. H., Ertelt, J. M., Xin, L. & Way, S. S. Pregnancy imprints regulatory memory that sustains anergy to fetal antigen. *Nature* **490**, 102–6 (2012).
36. Rosenblum, M. D., Way, S. S. & Abbas, A. K. Regulatory T cell memory. *Nat. Rev. Immunol.* **16**, 90–101 (2016).
37. Porrett, P. M. Biologic mechanisms and clinical consequences of pregnancy alloimmunization. *Am. J. Transplant* **18**, 1059–1067 (2018).
38. Berry, A. S. F. et al. Remodeling of the maternal gut microbiome during pregnancy is shaped by parity. *Microbiome* **9**, 146 (2021).
39. Bayar, E., Bennett, P. R., Chan, D., Sykes, L. & MacIntyre, D. A. The pregnancy microbiome and preterm birth. *Semin. Immunopathol.* **42**, 487–499 (2020).
40. Abalos, E. et al. Duration of spontaneous labour in ‘low-risk’ women with ‘normal’ perinatal outcomes: A systematic review. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **223**, 123–132 (2018).
41. Roos, N., Sahlin, L., Ekman-Ordeberg, G., Kieler, H. & Stephansson, O. Maternal risk factors for postterm pregnancy and cesarean delivery following labor induction. *Acta Obstet. Gynecol. Scand.* **89**, 1003–1010 (2010).

42. Fischer-Rasmussen, W. & Aegidius, J. Plasma oestriol in prolonged pregnancy. *Acta Obstet. Gynecol. Scand.* **51**, 25–30 (1972).
43. Schock, H. et al. Hormone concentrations throughout uncomplicated pregnancies: a longitudinal study. *BMC Pregnancy Childbirth* **16**, 146 (2016).
44. Romero, R. 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* **2**, 18 (2014).
45. Osman, I. et al. Leukocyte density and pro-inflammatory cytokine expression in human fetal membranes, decidua, cervix and myometrium before and during labour at term. *Mol. Hum. Reprod.* **9**, 41–5 (2003).
46. Yellon, S. M. Immunobiology of Cervix Ripening. *Front. Immunol.* **10**, 3156 (2019).
47. Malmström, E. et al. The importance of fibroblasts in remodelling of the human uterine cervix during pregnancy and parturition. *Mol. Hum. Reprod.* **13**, 333–41 (2007).
48. Witkin, S. S. et al. Influence of vaginal bacteria and D- and L-lactic acid isomers on vaginal extracellular matrix metalloproteinase inducer: implications for protection against upper genital tract infections. *MBio* **4**, (2013).
49. Beghini, J., Linhares, I. M., Giraldo, P. C., Ledger, W. J. & Witkin, S. S. Differential expression of lactic acid isomers, extracellular matrix metalloproteinase inducer, and matrix metalloproteinase-8 in vaginal fluid from women with vaginal disorders. *BJOG* **122**, 1580–5 (2015).
50. Walsh, S. W. et al. Protease Amplification of the Inflammatory Response Induced by Commensal Bacteria: Implications for Racial Disparity in Term and Preterm Birth. *Reprod. Sci.* **27**, 246–259 (2020).
51. Caughey, A. B., Nicholson, J. M., Cheng, Y. W., Lyell, D. J. & Washington, A. E. Induction of labor and cesarean delivery by gestational age. *Am. J. Obstet. Gynecol.* **195**, 700–705 (2006).
52. Caughey, A. B., Stotland, N. E., Washington, A. E. & Escobar, G. J. Who is at risk for prolonged and postterm pregnancy? *Am. J. Obstet. Gynecol.* **200**, 683.e1–5 (2009).
53. Kolkman, D. G. E. et al. The Bishop score as a predictor of labor induction success: a systematic review. *Am. J. Perinatol.* **30**, 625–30 (2013).
54. Crane, J. M. G. Factors predicting labor induction success: a critical analysis. *Clin. Obstet. Gynecol.* **49**, 573–84 (2006).
55. Serrano, M. G. et al. Racial/ethnic diversity in the dynamics of the vaginal microbiome during pregnancy. *Nat. Med.* **25**, 1001–1011 (2019).
56. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
57. Virtanen, S. et al. Metagenome-validated Parallel Amplicon Sequencing and Text Mining-based Annotations for Simultaneous Profiling of Bacteria and Fungi: Vaginal Microbiome and Mycobiota in Healthy Women. *Res. Sq.* (2021) doi:10.21203/rs.3.rs-321778/v1.

58. Oksanen, J. et al. vegan: Community Ecology Package. R package version 2.5-6. <https://cran.r-project.org/package=vegan> (2019).
59. Korpela, K. mare: Microbiota Analysis in R Easily. R package version 1.0. <https://github.com/katrikorpela/mare> (2016).
60. R Core Team. R: A language and environment for statistical computing. <https://www.r-project.org/> (2020).
61. Venables, W. N. & Ripley, B. D. *Modern Applied Statistics with S*. (Springer, New York, 2002).
62. Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & Team, R. C. \_nlme: Linear and Nonlinear Mixed Effects Models\_. R package version 3.1-149. <https://cran.r-project.org/package=nlme> (2020).
63. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).

## Table

Table 1. Characteristics of the study population.

Characteristics	All	Parity		p-value	Gestational age		p-value
		Nulliparous	Multiparous		<41.0	≥41.0	
Number of women	324	188	136		155	169	
Age, years							
Mean (SD, range) <sup>1</sup>	32.1 (4.8, 20-47)	31.6 (4.7, 20-43)	32.8 (4.8, 21-47)	0.02	31.9 (4.7, 20-47)	32.2 (4.9, 20-43)	0.51
BMI, kg/m <sup>2</sup>							
Median (IQR) <sup>2</sup>	23.1 (21.2-26.4)	23.1 (21.3-26.5)	23.1 (21.2-26.0)	0.91	23.0 (21.0-25.0)	24.0 (21.5-27.3)	0.02
Nullipara <sup>3</sup>	188/324 (58.0)				82/155 (52.9)	(62.7)	0.07
Primigravida	139/188 (73.9)				60/82 (73.2)	(74.5)	0.83
Nulliparous multigravida	49/188 (26.1)				22/82 (26.8)	(25.5)	
Current smoking <sup>3</sup>	10/324 (3.1)	2/188 (1.1)	8/136 (5.9)	0.01	7/155 (4.5)	3/169 (1.8)	0.20
Assisted reproductive technology in current pregnancy <sup>3</sup>	22/324 (6.8)	17/188 (9.0)	5/136 (3.7)	0.06	16/155 (10.3)	6/169 (3.6)	0.02
Use of probiotics*, <sup>3</sup>	192/303 (63.4)	107/171 (62.6)	85/132 (64.4)	0.74	88/141 (62.4)	(64.2)	0.75
Antibiotic use (<3 months)*, <sup>3</sup>	45/302 (14.9)	23/172 (13.4)	22/130 (16.9)	0.39	16/140 (11.4)	(17.9)	0.12
Intercourse <48h*, <sup>3</sup>	46/304 (15.1)	26/174 (14.9)	20/130 (15.4)	0.92	15/142 (10.6)	(19.1)	0.04
High education (tertiary degree)*, <sup>3</sup>	198/303 (65.3)	117/171 (68.4)	81/132 (61.4)	0.20	107/164 (65.2)	91/139 (65.5)	0.97
Married or cohabiting*, <sup>3</sup>	288/304 (94.7)	164/172 (95.3)	124/132 (93.9)	0.59	133/139 (95.7)	155/165 (93.9)	0.50
<b>Pregnancy related characteristics</b>							
Gestational age at sampling, weeks							
Median (IQR) <sup>2</sup>	41.0 (39.7-41.7)	41.4 (40.1-41.7)	40.6 (39.4-41.7)	0.03	39.7 (39.1-40.3)	41.7 (41.6-41.7)	<0.001
≥H41+0 <sup>3</sup>	169/324 (52.2)	106/188 (56.4)	63/136 (46.3)	0.07			
Gestational age at birth, weeks							
Median (IQR) <sup>2</sup>	41.2 (39.7-41.9)	41.6 (40.1-42.0)	40.6 (39.4-41.9)	0.001	39.7 (39.1-40.4)	41.9 (41.9-42.0)	<0.001
Contractions at sampling <sup>3</sup>	103/324 (31.8)	63/188 (33.5)	40/136 (29.4)	0.43	95/155 (61.3)	8/169 (4.7)	<0.001
Gestational diabetes <sup>3</sup>	79/324 (24.4)	46/188 (24.5)	33/136 (24.3)	0.97	37/155 (23.9)	(24.9)	0.84
Maternal postpartum infection†, <sup>3</sup>	32/324 (9.9)	18/188 (9.6)	14/136 (10.3)	0.83	11/155 (7.1)	(12.4)	0.11
Birthweight, g						3808	
Mean (SD, range) <sup>1</sup>	3679 (474, 2636-5130)	3627 (445, 2690-5020)	3751 (502, 2636-5130)	0.02	3538 (421, 2636-4744)	(2740-5130)	<0.001

BMI, body mass index; SD, standard deviation; IQR, interquartile range

Data are n (%) unless otherwise specified.

\*Missing information for data collected from questionnaires, total number of women for which data is known is shown for each variable.

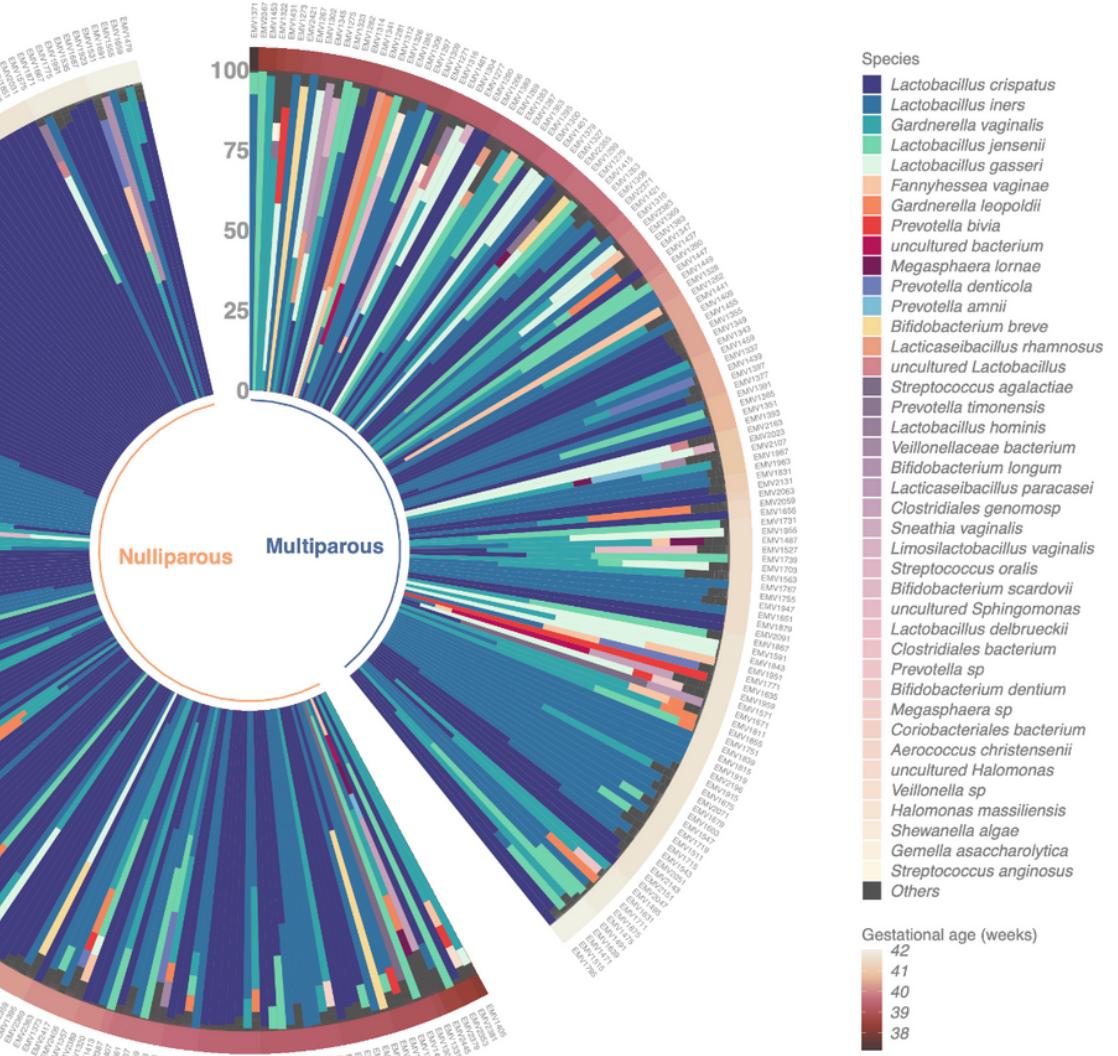
<sup>†</sup>Postpartum infections included endometritis n=24, episiotomy infection n=6, urinary tract infection n=1, post-cesarean section wound infection n=1.

1 t test

## <sup>2</sup> Mann-Whitney U-test

<sup>3</sup> Chi square and Fisher's exact test when appropriate

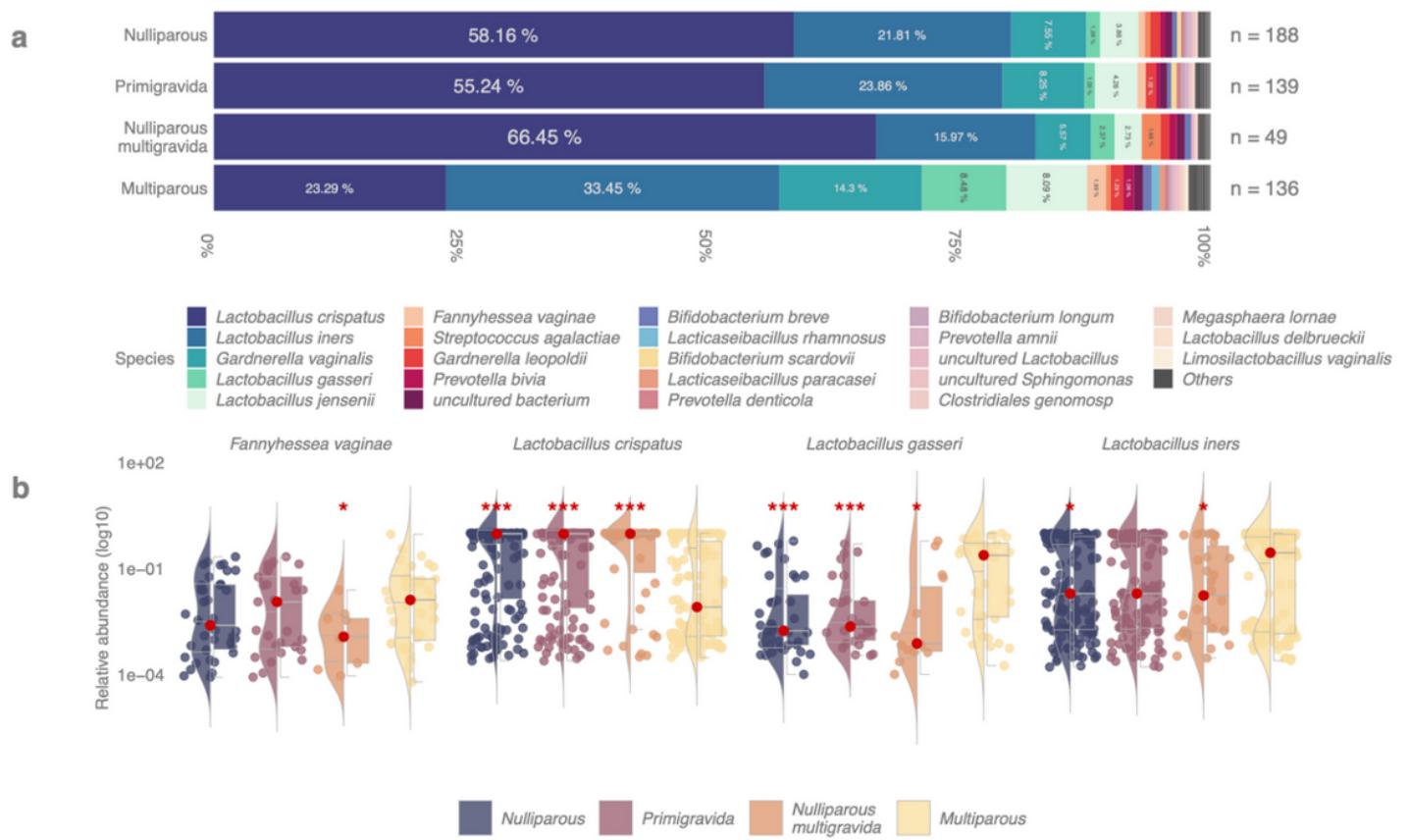
# Figures



**Figure 1**

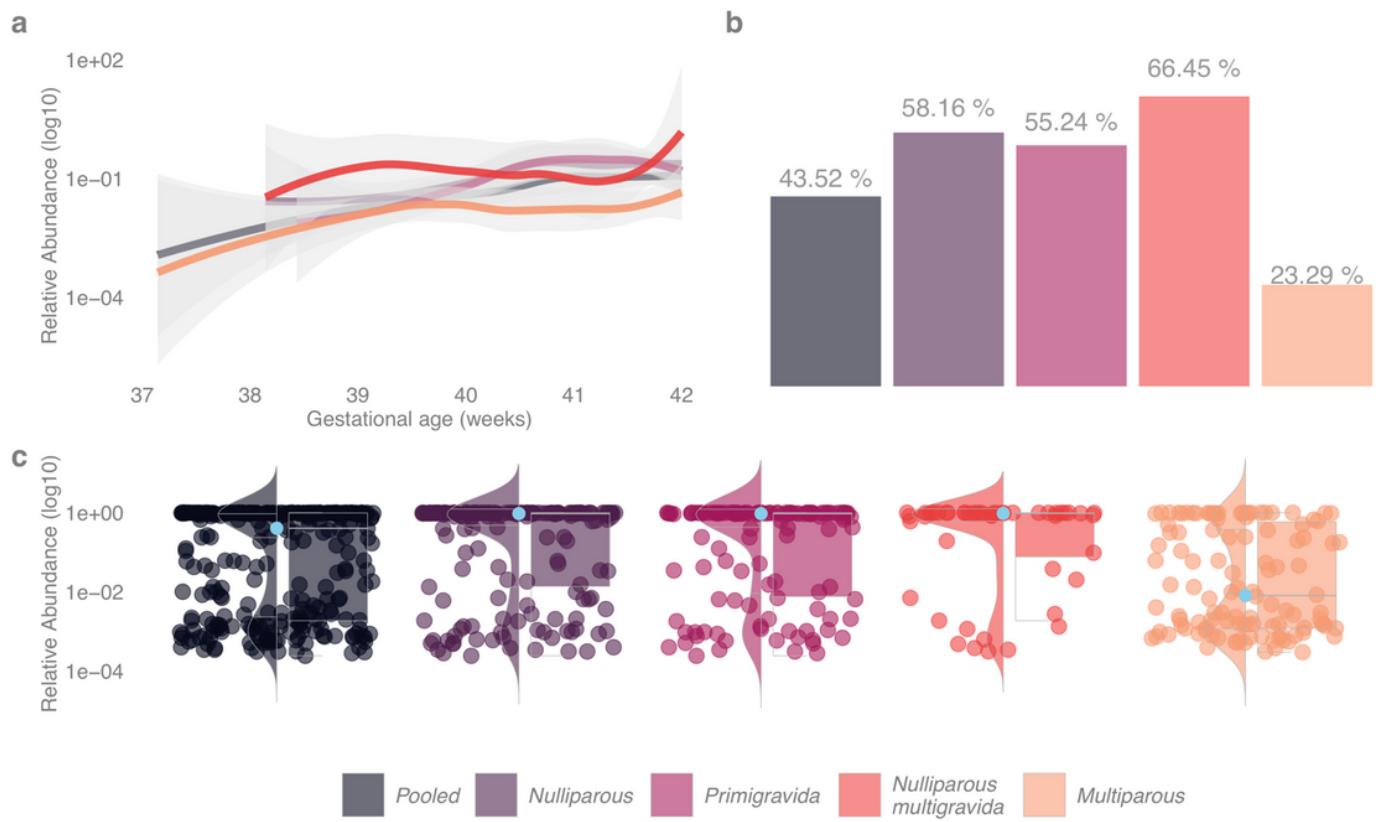
Relative abundances of bacterial taxa across all 324 samples. Circular stacked bar chart showing the bacterial relative abundances of all the samples in the study cohort (n=324), divided based on parity – no previous pregnancies or deliveries (nulliparous) (n=188) and one or more previous deliveries (multiparous) (n=136). The samples have been ordered based on two criteria 1) gestational age in weeks,

increasing clockwise (outer band) and 2) within each gestational age bracket, whether one of the top 3 bacteria were the most abundant (>50% composition), ordered clockwise from None, to Gardnerella vaginalis, Lactobacillus iners, and Lactobacillus crispatus dominant samples.



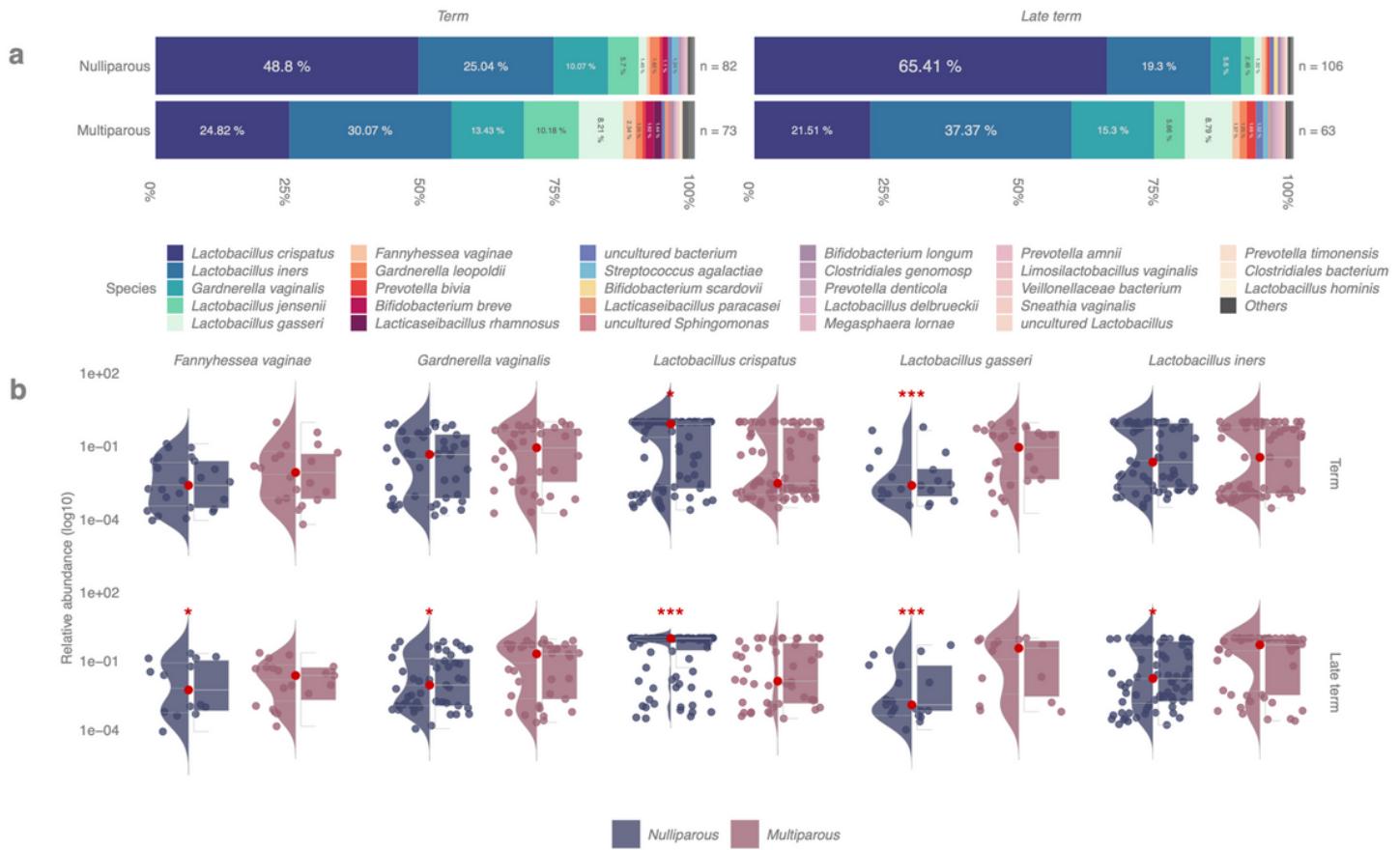
**Figure 2**

Abundance of bacterial taxa in pregnant women based on previous pregnancy history. a Stacked bar plot depicting the mean bacterial relative abundances of samples from women without prior pregnancies or deliveries (nulliparous) (n=188), women without prior pregnancies (primigravida) (n=139), women with prior spontaneous or induced abortion (nulliparous multigravida) (n=49), and women with previous deliveries (multiparous) (n=136). b Violin + box + jitter plots showing the distribution of taxa ( $\log_{10}$  relative abundance) with significantly different abundances between the sample groups. The whiskers on the boxplot represent the 1.5 interquartile range and the median value is shown as a red dot. The horizontal lines on the violin plots represent the 25th, 50th, and 75th quantiles. Each point on the jitter plot represents a sample, highlighting the density and frequency of occurrence of a taxon. Asterisks indicate whether there were statistically significant differences between the subgroups compared to the multiparous samples, \*\*\* q <0.001, \*\* q <0.01, \* q <0.05.



**Figure 3**

*Lactobacillus crispatus* distribution in relation to gestational age and parity. Summary of *L. crispatus* distribution across all samples (pooled, n=324), women with no previous pregnancies or deliveries (nulliparous) (n=188), women without prior pregnancies (primigravida) (n=139), women with prior spontaneous or induced abortion (nulliparous multigravida) (n=49), and women with one or more previous deliveries (multiparous) (n=136). a Local regression models representing the change in relative abundance (log<sub>10</sub> scale) of *L. crispatus* with increasing gestational age in weeks. b Bar plot showing the mean relative abundance of *L. crispatus* in each subgroup across all gestational weeks. c Violin + box + jitter plots showing the distribution of taxa (log<sub>10</sub> relative abundance) with significantly different abundances between the sample groups. The whiskers on the boxplot represent the 1.5 interquartile range and the median value is shown as a sky-blue dot. The horizontal lines on the violin plots represent the 25th, 50th, and 75th quantiles. Each point/dash on the jitter plot represents a sample, highlighting the density and frequency of occurrence of *L. crispatus*.



**Figure 4**

Relative abundances of bacterial taxa based on gestational age and parity. a Stacked bar plot depicting the mean bacterial relative abundances of samples from women whose duration of pregnancy was <41.0 gestation weeks (Term) (n=155) or >41 weeks (Late term) (n=169). The graphic is further split based on parity - no previous deliveries (nulliparous) (n=188) and one or more previous deliveries (multiparous) (n=136). b Violin + box + jitter plots showing the distribution of taxa (log10 relative abundance) with significantly different abundances between nulli- and multiparous in term and late term groups. The whiskers on the boxplot represent the 1.5 interquartile range and the median value is shown as a red dot. The horizontal lines on the violin plots represent the 25th, 50th, and 75th quantiles. Each point on the jitter plot represents a sample, highlighting the density and frequency of occurrence of a taxon. Asterisks indicate whether there were statistically significant differences between the subgroups compared to the multiparous samples, \*\*\* q <0.001, \*\* q <0.01, \* q <0.05.

## Supplementary Files

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

- Descriptionofadditionalsupplementaryfiles.docx
- SupplementaryInformation.pdf

- SupplementaryData1.xlsx