Vaginal microbiota and cytokine levels predict preterm delivery in Asian women

Manoj Kumar
Sidra Medical and Research Center
Selvasankar Murugesan
Sidra Medical and Research Center
Parul Singh
Sidra Medical and Research Center
Marwa Saadaoui
Sidra Medical and Research Center
Duaa Ahmed Elhag
Sidra Medical and Research Center
Annalisa Terranegra
Sidra Medical and Research Center
Basirudeen Syed Ahamed Kabeer
Sidra Medical and Research Center
Alexandra K Marr
Sidra Medical and Research Center
Tomoshige Kino
Sidra Medical and Research Center
Tobias Brummaier
Mahidol University Faculty of Tropical Medicine
Rose McGready
Mahidol University Faculty of Tropical Medicine
François Nosten
Mahidol University Faculty of Tropical Medicine
Damien Chaussabel
Sidra Medical and Research Center
Souhaila Al Khodor (salkhodor@sidra.org)
Sidra Medical and Research Center https://orcid.org/0000-0003-4858-7130

Research

Keywords: microbiota, metagenomics, 16S rRNA gene sequencing, dysbiosis, vaginal cytokines, Nugent scoring, Asian

Posted Date: August 17th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-56673/v1

Vaginal microbiota and cytokine levels predict preterm delivery in Asian women

Manoj Kumar[1], Selvasankar Murugesan[1], Parul Singh[1], Marwa Saadaoui[1], Duaa Ahmed Elhag[1], Annalisa Terranegra[1], Basirudeen Syed Ahamed Kabeer[1], Alexandra K Marr[1], Tomoshige Kino[1], Tobias Brummaier[2-5], Rose McGready[2,3], François Nosten[2,3], Damien Chaussabel[1] and Souhaila Al Khodor[1]*

[1]Research Department, Sidra Medicine, Doha, Qatar
[2]Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Mae Sot, Thailand
[3]Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
[4]Swiss Tropical and Public Health Institute, Basel, Switzerland
[5]University of Basel, Basel, Switzerland

***Corresponding author:** Dr. Souhaila Al Khodor (salkhodor@sidra.org)

Running title: Preterm birth in Asian women: Role of the Vaginal Microbiome

Abstract

Background: Preterm birth (PTB) is the most common cause of neonatal morbidity and mortality worldwide. Approximately half of PTBs are linked with microbial etiologies, including pathologic changes to the vaginal microbiota, which vary according to ethnicity. Globally more than 50% of PTBs occur in Asia, but studies of the vaginal microbiome and its association with pregnancy outcomes in Asian women are lacking. This study aimed to characterize the vaginal microbiome and cytokine environment of 18 Karen and Burman pregnant women who delivered preterm and 36 matched controls delivering at full term.

Results: Using 16S ribosomal RNA gene sequencing we identified a predictive vaginal microbiota signature for PTB that was detectable as early as the first trimester of pregnancy, characterized by higher levels of *Prevotella buccalis*, and lower levels of *Lactobacillus crispatus and Finegoldia*, accompanied by decreased levels of cytokines including IFNγ, IL-4 and TNFα.

Conclusion: Our findings highlight new opportunities to predict PTB in Asian women in low-resource settings who are at highest risk of adverse outcomes from unexpected PTB, as well as in Burman/Karen ethnic minority groups in high-resource regions.

Keywords: microbiota, metagenomics, 16S rRNA gene sequencing, dysbiosis, vaginal cytokines, Nugent scoring, Asian

Background

Preterm birth (PTB), defined as birth before 37 weeks of gestation, is the leading worldwide cause of neonatal mortality, and of morbidity and mortality in children under five in high income countries [1]. Nearly 15 million pregnant women experience PTB every year worldwide, representing around 12% of all deliveries [1, 2], and resulting in approximately 1.1 million infant deaths annually [3]. Children born prematurely face an increased risk of complications attributed to multiple organ immaturity, as well as possible lifelong effects on neurocognitive development, visual disorders, and increased risk of chronic disease [4]. Although PTB is a global public health concern, certain racial and ethnic groups are more predisposed to PTB than others [2, 5, 6]. Altogether, more than half of PTBs occur in Asia [6], but this ethnically-diverse population group is relatively under-studied in terms of PTB causation/risk factors. What is clear is that the combination of environmental influences, high prevalence of infectious disease, and low healthcare resource settings places these women at high risk of adverse pregnancy outcomes following PTB. There is an urgent need for clinically applicable strategies to identify those Asian women in the community at a high risk of PTB. Alongside, knowledge of specific PTB predictive factors for Asian minority groups in other countries will facilitate appropriate medical care in these higher-resource settings.

Pregnancy is an important “formative period” where a series of interconnected physiological and cellular processes aim to support healthy fetal development [7]. These processes include maternal and paternal genetic factors, hormonal changes, immune system modulation, environmental factors, the microbiome and others [8, 9]. While we have known for some time that microbial factors could underpin as many as 50% of all PTBs [10, 11], only recently has the association between specific changes in the vaginal microbiome and pregnancy complications started to be unraveled [12-14]. During pregnancy, the composition of the vaginal microbiome undergoes an overall change in microbial diversity and exhibits clade-specific enrichments and depletions [12-16] that vary according to ethnicity. For example, in Caucasian women, adverse pregnancy outcomes including PTB have been linked with a shift from a *Lactobacilli*-rich vaginal microbiome to a more complex microbial community of *Gardnerella, Prevotella,* and *Lachnospiraceae* family members (bacterial vaginosis (BV)-associated bacterium-I (BVAB-I) [14, 16, 17]; while African women are less likely

to have a vaginal microbiome dominated by *Lactobacillus* species [14, 18]. Despite the high burden of PTB in Asian regions, few studies have addressed the composition of the vaginal microbiome in reproductive age women living in this geographical area [19-21] but not in the context of PTB. Moreover, the high ethnic diversity of the region means that some populations are completely unstudied. This is especially important in areas with low healthcare resources, where an unexpected PTB can have devastating consequences for mother and baby. Therefore, population-specific studies are needed to improve our knowledge of the vaginal microbiome composition during pregnancy and its association with PTB in Asian communities.

Here we report the results from a study of 54 Karen and Burman ethnicity pregnant women recruited prospectively at 8-14 weeks gestation at the Shoklo Malaria Research Unit (SMRU), Thailand, as part of the larger Molecular Signature in Pregnancy study [22] (Table 1 and Supplementary Table 1 for cohort characteristics). Following routine antenatal care procedures, vaginal samples were collected at the time of enrollment (trimester 1, T-1), trimesters 2 and 3 (T-2, T-3), and at delivery (Fig. 1). As a start, we compared the vaginal microbiome composition between Karen and Burman pregnant women. Despite being genetically different [23], we found that the vaginal microbiome composition in both ethnicities is similar and hence they were grouped together in all the following analyses (Supplementary Fig. 1). We then compared the vaginal microbiome composition and vaginal cytokine levels of women who experienced PTB, defined as delivery before 37 weeks (n=18), and controls who had a full term birth (TB) at or after 37 weeks (n=36) and were case-matched matched by age, parity, and gravida at a 1:2 ratio. There were no significant differences in maternal height, weight, body mass index or delivery mode between PTB and TB groups (Table 1 and Supplementary Table 1). The mean gestational age at delivery for those who delivered PTB and TB was 36.2 and 39.5 weeks respectively, and as expected, preterm neonates had a lower birth weight compared to term neonates (Table 1). For each of the 633 samples collected from the 54 participants, we: used Gram staining to calculate the Nugent score and thereby identify bacterial vaginosis; generated a bacterial profile of the vaginal environment using 16S rRNA gene sequencing; and measured levels of a panel of cytokines. Combining these features, we went on to generate a predictive vaginal microbial signature that is observed as early as the first trimester of pregnancy and identified a correlative cytokine profile for PTB in these groups of Asian women.

Methods

Study design

This observational, prospective, pregnancy-delivery-postpartum cohort enrolled pregnant women with an unremarkable medical and obstetric history and was a collaboration between Sidra Medicine, Doha, Qatar, and SMRU, Mae Sot, Thailand [22]. SMRU is a field station of the Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand, and is part of the Mahidol-Oxford Research Unit, which combines research and humanitarian work to serve rural and disadvantaged migrant and refugee populations on the Thailand-Myanmar border. The study was conducted in accordance with the Declaration of Helsinki and followed ICH Guidelines for Good Clinical Practice. All study procedures were reviewed and approved by the Ethics Committee, Faculty of Tropical Medicine, Mahidol University, Thailand (TMEC 15-062), the Oxford Tropical Research Ethics Committe, University of Oxford, UK (OxTREC: 33-15) and Sidra Medicine, Qatar (IRB protocol #1705010909). Written informed consent or consent by thumbprint, in the case of illiterate participants, was obtained prior to study enrolment. The study was registered at clinicaltrial.gov (registration number NCT02797327).

Participant recruitment, clinical history, and sample collection

First trimester pregnant women with a viable, singleton pregnancy were enrolled at SMRU's antenatal care (ANC) clinics on the Thailand-Myanmar border. Gestational age was determined by early ultrasound scan and women with an estimated gestational age from 8 weeks 0 days to 13 weeks 6 days (T1) were eligible. At the time of recruitment, comprehensive maternal demographic information, medical, and obstetric history were recorded together with a detailed physical and obstetric examination. Women were followed-up each trimester (T2: 20-24 weeks, T3: 32-35 weeks) and during delivery.

Vaginal swab samples were collected from the posterior formix at a sampling point in each trimester, and at delivery, by a trained midwife using the Copan Eswab™ collection system. Three swabs were taken at each timepoint: one to extract genomic bacterial DNA, one to prepare a Gram stain smear to assess the Nugent score, and one for measurement of vaginal cytokines. Samples were transferred daily from the clinic sites to the central laboratory facility and stored at -80 degrees Celsius. For

international shipment samples were kept on dry ice in Styrofoam boxes that were equipped with temperature monitors.

Here we present results from 18 preterm cases of singleton, non-medically indicated PTB women who delivered between 32 weeks and 2 days of gestation and 36 weeks 6 days of gestation and were selected from a cohort of 428 women. We case matched the preterm participants 1:2 with participants who delivered at term (≥37 weeks) with a singleton pregnancy. The case matching was performed based on age, parity and gravida [24].

Nugent scoring

Vaginal swab smears were screened for BV using the Nugent scoring method described earlier [25]. Briefly, frozen vaginal swab samples were thawed on ice, vortexed vigorously for 1 min then swabs were rolled onto slides, air dried and Gram-staining using the standard protocols. Vaginal smears were evaluated by the Nugent scoring method for the relative abundance of three types of bacterial cell morphotypes: large Gram-positive rods (*Lactobacillus* morphotypes), small Gram-negative rods and cocci (*Gardnerella vaginalis*, Bacteroides), and curved Gram-negative bacilli [25]. The Nugent scores range from 0–3 (normal), 4–6 (intermediate), and 7– 10 (indicative of BV) [26].

DNA extraction and 16S rRNA gene sequencing

The total DNA from vaginal swabs was extracted using the modified protocol MoBio Powersoil modified Method #3 as previously published [27]. DNA concentration was measured using Nanodrop.

The V1-V3 regions of the 16S rDNA were amplified using forward primers: 27F with 12 bp golay barcodes containing a specific Illumina 5' adapter for each sample and a common reverse primer 515 R [27]. In brief, PCR was performed in triplicate in a 50 µL reaction mixture containing 10 ng of template DNA and 2x Phusion HotStart Ready Mix. The following thermal cycling conditions were used: 5 min of initial denaturation at 94 °C; 25 cycles of denaturation at 94 °C for 30 s, annealing at 62 °C for 30 s, and elongation at 72 °C for 30 s; and the last step at 72 °C for 10 min. The amplified PCR products of approximately 650 bp in size from each sample were pooled in equimolar concentrations. This pooled PCR product was purified using AgenCourt AMPure XP magnetic beads. High throughput sequencing was performed

on an Illumina MiSeq 2 × 300 platform (Illumina, Inc. San Diego) in accordance with the manufacturer's instructions. Image analysis and base calling were carried out directly on the MiSeq.

Vaginal 16S rRNA taxonomic profiling

Sequenced data were demultiplexed using MiSeq Control Software (MCS) then quality controlled using FastQC [28]. Forward and reverse end sequences of respective samples were merged through the PEAR tool [29] and sequence reads of quality score < 20 were discarded. All merged reads were trimmed to 160bp>Reads<500bp using the Trimmomatic tool [30]. Trimmed FASTQ files were converted into FASTA files. Demultiplexed FASTA files were analyzed using QIIME (Quantitative Insights Into Microbial Ecology) v1.9.0 pipeline [31]. Operational taxonomic units (OTUs) were generated by aligning against the SILVA database.

Microbial diversity analysis

Alpha diversity was measured by R software, using the phyloseq package [32]. Beta diversity was represented using Phylogenetic beta diversity metrics [33] and the differences in the beta diversity were presented as principal coordinate analysis using the QIIME (Quantitative Insights Into Microbial Ecology) v1.9.0 pipeline. We then performed a longitudinal analysis to evaluate the relationship between vaginal microbial community, delivery status and the stage of pregnancy using the ggplot package in the R software.

Vaginal cytokine profiling

Cytokine levels in the vaginal swab samples were quantified using Bio-Red Pro cytokine 8-Plex assay kit with a Luminex 3D system. In addition, a single-Plex assay was used for IL-1B (Bid-Red) [34]. Frozen vaginal swab samples were thawed, vortexed vigorously for 1 min and the swab squeezed on the wall of the tube to maximize elution. The swab samples were then centrifuged at 700 xg for 10 minutes. The Bio-Rad assay was performed using samples and serial dilutions of standards in duplicate, according to the manufacturer's instructions. Cytokine levels were analyzed using LuminoXponent software.

Integrative correlation analysis of cytokine levels and vaginal microbial taxa

To explore the correlation between differentially-present vaginal microbiota and cytokine levels at the time of delivery, a comprehensive integrative sCCA was performed as described earlier [14, 35]. Briefly, an integrative analysis of log-transformed cytokine data and 16S rRNA vaginal taxonomic data was performed to explore the correlation between two data sets of quantitative variables measured on the same subjects using sCCA . The most abundant bacterial taxa were designated as present if they comprised $\geq$0.1% of the total vagitype profile in either group, and nine cytokines were selected for the analysis. For TB and PTB, sCCA was performed separately using the sgcca package in R and displayed in a correlation circle plot [36]. All variables with a strong positive correlation are grouped together, while variables with negative correlations are plotted opposite each other.

Statistical analysis

Statistical significance of alpha diversity measures such as Observed, Chao1, Shannon and Simpson indices were calculated using minitab17 (Minitab statistical software). *P*-values lower than 0.05 were considered statistically significant. Adonis was used to calculate the distance matrix difference between the categories included in this study using unweighted beta diversity parameters [33]. Association between cytokines profile and vaginal microbiome data was performed using CCA as described earlier [35].

For the cytokine data analysis, all cytokine values were log-transformed before analysis to achieve normality and homogeneity of the raw values. The unpaired *t*-test analysis with Welch's correction was used to determine significant differences between cytokine levels of women who experienced PTB versus TB. Statistical analyses were performed using GraphPad Prism 8 (GraphPad Softwares Inc. USA).

Results

PTB is positively associated with bacterial vaginosis in Burman and Karen women

In Caucasian women a healthy vaginal microbiome is generally dominated by *Lactobacillus species* such as *L. crispatus* [13, 18]; while an imbalanced microbiome (termed bacterial vaginosis (BV)) is characterized by a decrease in *Lactobacillus species* and an increase in mixed anaerobes such as *Gardnerella*, *Atopobium*, *Prevotella, Megasphaera* species, and others [37, 38]. As it is unknown whether the same species characterize vaginal microbial balance/imbalance in this group of Asian women, we began by identifying and comparing the bacteria present in vaginal swab smears from our TB and PTB cohorts. We used Nugent scoring, which is based on microscopic morphotype enumeration of Gram-positive *Lactobacilli* vs. Gram-negative bacteria, to define normal (score = 0-3), altered/intermediate (score = 4-6) or BV (score = 7-10) categories [26]. Representative images for each category are shown in Fig. 2a. Women in the PTB group were more likely to exhibit either high or low Nugent scores during pregnancy than women in the TB group, of which the majority exhibited Nugent scores in the intermediate category (Fig. 2b, c and Supplementary Fig. 2). However, by delivery 12 of the 18 PTB women exhibited BV, leading to a significantly higher average Nugent score compared to those in the TB group (Fig. 2c, fourth panel). This is consistent with the previously-reported positive association between PTB and BV in a predominantly Caucasian cohort of women [39], but is the first evidence of a similar association in an Asian cohort.

High microbial diversity characterizes the vaginal microbiome in PTB

Having established that increased bacterial diversity/BV was positively associated with PTB in our cohort, we next asked which bacterial species were involved. To assess the detailed changes in microbiome composition we generated vaginal microbiota relative abundance profiles using 16S rRNA taxonomic analysis, and saw clear differences between TB and PTB groups at both the phylum and species levels (Supplementary Fig. 3 and Fig. 3).

Given the apparent variability in distribution of different microbial species between TB and PTB groups, we next applied diversity analysis to understand whether the variability itself was important. We applied two diversity analyses: alpha diversity

(Fig. 3b) measures the average species diversity within a sample community and was calculated first from the total number of unique operational taxonomic units (OTUs) observed per sample; then using the Chao1 abundance-based richness estimator, which is sensitive to rare OTUs [40]; and lastly by Shannon [41] and inverse Simpson (InvSimpson) [42] approaches which are more dependent on highly-abundant OTUs and less sensitive to rare OTUs [32]. Following alpha analysis we then employed beta diversity to understand the divergence in community composition between samples (Fig. 3c), which we assessed using principal coordinate analysis (PCoA) based on Bray-Curtis dissimilarities [33]. When we compared the overall vaginal microbial richness at delivery by alpha diversity analysis, we saw that women who delivered preterm showed significantly more microbial richness (Observed, $p=0.005$; Chao1, $p=0.005$); coupled with greater beta diversity ($p=0.001$) compared to women who had term deliveries (Fig. 3b, c).

We then clustered vaginal microbial diversity using Euclidean distance matrices with Ward linkage (Fig. 4) into the following community state types (CST), as previously described by Ravel *et al*: CST-I (*Lactobacillus crispatus*-dominated), CST-II (*Lactobacillus gasseri*-dominated), CST-III (*Lactobacillus iners*-dominated), CST-IVA (lower abundance of *Lactobacillus spp*, together with low proportions of various anaerobic bacteria such as *Anaerococcus*, *Corynebacterium*, *Finegoldia*, or *Streptococcus*); CST-IVB (higher abundance of the genera *Atopobium*, *Prevotella*, *Parvimonas*, *Sneathia*, *Gardnerella*, *Mobiluncus*, or *Peptoniphilus* and several other taxa often associated with high Nugent scores); and CST-V (*Lactobacillus jensenii*-dominated) [18, 43]. The most commonly observed was CTS III (*L. iners*) followed by CST I (*L. crispatus*) and CST IVB (dominated by *Prevotella Buccalis*) respectively (Fig. 4). Comparing TB and PTB groups confirmed substantial differences in the overall microbial profiles, with TB women typically exhibiting a higher prevalence of *Lactobacillus*-dominated vagitypes (CST-I, and -III), and PTB women exhibiting a significantly higher frequency of CST-IV ($p<0.0001$) (Fig. 5a). We then looked at the dynamics of the vaginal microbiome during pregnancy to assess whether the vaginal microbial communities in TB or PTB groups persist across the sampling points or whether there is a transition between different vagitypes during pregnancy. We observed that most women in the TB group had a vaginal environment dominated by either CST-III (50%) or CST-I (36%) throughout the pregnancy period (Fig. 5a), while, women in the PTB group had a higher prevalence of CST-IV vagitype (40%) as early

as the first trimester (Fig. 5a). The profiles of CSTs for each pregnant woman as a function of gestational time clearly highlight the enrichment of CST-IVB in women in the PTB group (Fig. 5b). Similar associations between higher levels of CST-IV and lower levels of *Lactobacillus species* and PTB were also observed in a predominantly African cohort of women [14], highlighting that their abundance might have some conserved influence on pregnancy outcomes.

Taken together, these data show that the vaginal microbiome associated with TB in this Asian cohort is generally dominated by CST-III throughout the course of pregnancy. In contrast, PTB is significantly associated with CST-IVB, characterized by a lower abundance of *Lactobacillus species* and a higher abundance of mixed anaerobic bacteria known to be associated with high Nugent scores [25, 37], as well as with greater overall microbial diversity across the PTB group. These effects were present throughout pregnancy starting from the first trimester.

Low *L. crispatus,* low *Finegoldia* and high *P. buccalis* precede PTB in Burman and Karen women

Given the high diversity of microbes in samples from women who experienced PTB, we next asked whether any microbial species were particularly associated with premature delivery. We first compared the relative levels of the eleven most abundant bacterial taxa (Fig. 4) between the TB and PTB groups (Fig. 6). Of these, three were significantly different: in PTB women, *P. buccalis* was significantly more abundant throughout pregnancy from as early as the first trimester, compared to women with TB (Fig. 6a-d); while *Finegoldia* and *L. crispatus* were significantly less abundant during the pregnancies of PTB women compared to TB women, also as early as the first trimester (Fig. 6a-d). These data are similar to earlier findings in women from African ancestry showing that lower levels of *L. crispatus* in early and mid-pregnancy are associated with PTB [14].

We further confirmed the importance of longitudinal trends in abundance of the three most significant bacteria using ggplots incorporating delivery status (TB or PTB), longitudinal microbial abundance and timepoint during pregnancy (Fig. 6e). We observed that women in the PTB group exhibited a significant increase in relative abundance of *P. buccalis* ($P<0.0001$), whereas women who delivered at term showed a significant increase in *L. crispatus* ($P=0.0131$), and *Finegoldia* ($P<0.0001$) from trimester 1 and throughout pregnancy (Fig. 5e). In summary, we have identified early

trends in abundance of specific bacterial taxa that occur during pregnancies that end in PTB, and distinct trends that characterize TB pregnancies: these trends are evident as early as the first trimester, indicating their possible predictive potential.

Vaginal cytokines differ between PTB and TB pregnancies

The vaginal microbial environment is maintained by a delicate interplay between host factors (ethnicity, genetics and immune mediators) and microbial biology. While altered vaginal cytokine patterns have been described during pregnancy, their relationship to the vaginal microbiota and pregnancy outcomes is incompletely defined. In this study, we measured the vaginal levels of nine cytokines: interleukin (IL)-1β, IL-2, IL-4, IL-6, IL-8, IL-10, granulocyte-macrophage colony-stimulating factor (GM-CSF), tumor necrosis factor (TNF)-α, and interferon (IFN)-γ, and asked whether these levels differed between TB and PTB groups. In PTB pregnancies we found significantly higher levels of IL-8 in trimester 3 samples compared to TB pregnancies at the same timepoint; while by the point of delivery IL-8, IL-6 and IL-10 were all higher in the PTB group (Fig. 7). In contrast, women with PTB showed significantly lower levels of TNF-α and IL-2 in the first two trimesters than did TB women (Fig. 7). However, the most robust differences were seen in levels of IFN-γ and IL-4 which were significantly lower in the PTB group throughout their entire pregnancies and at delivery (Fig. 7). Thus, we have identified several potential biomarkers in the vaginas of PTB women that distinguish their pregnancies from those of TB women, from as early as the first trimester.

Integrative analysis of vaginal cytokine levels and microbial profiles

To bring together our data on the vaginal microbiome and cytokine environment linked with PTB in our cohort of Burman and Karen Asian women we performed an integrative sparse canonical correlation analysis (sCCA) which assessed the association between the abundance of the main microbial taxa, the levels of cytokines, and pregnancy outcomes. For each participant, the samples collected at the first trimester (8-14 weeks) were characterized. In the women who delivered at term, we observed a moderate positive correlation between *L. crispatus*, *L. gasseri* and *Finegoldia* as well as a strong negative correlation between *L. crispatus*, *L. gasseri*, *Finegoldia* and several taxa known to be associated with dysbiosis including *C.*

trachomatis, *P. buccalis, Prevotella, S amnii*, *S sanguinegens* (Fig. 8a). We also observed a strong negative correlation between *L. crispatus*, *L. gasseri*, *Finegoldia* and *L. iners* (Fig. 8a).

In contrast, in women who experienced PTB, the abundance of *L. crispatus* and *L. gasseri* was negatively correlated *with Finegoldia,* and similar negative correlation was observed between *L. crispatus*, *L. gasseri* with *C. trachomatis*, *P. buccalis* and *Prevotella* (Fig. 8b). Similar to the TB group, a strong negative correlation between *L. crispatus*, *L. gasseri*, and *L. iners* was observed (Fig. 8b). While *P. buccalis* and other taxa known to be associated with dysbiosis were negatively correlated with the cytokines tested, abundance of *Finegoldia* was strongly correlated with levels of IL-2 and IL-4 in the PTB group, and this was observed as early as the first trimester (Fig. 7b).

On the other hand, when we performed the sCCA integrative analysis at the time of delivery, a strong positive correlation was observed between pro-inflammatory cytokines such as IL-1b, IL-6 and IL- 8 with *P. buccalis* and *Prevotella 6*, but a strong negative correlation of the same cytokines with *L. crispatus* was observed (Supplementary Fig. 4).

In this study, we identified a predictive vaginal microbiota signature for PTB that was detectable as early as the first trimester of pregnancy, characterized by higher levels of *P. buccalis*, and lower levels of *L. crispatus and Finegoldia*, accompanied by decreased levels of cytokines including IFNγ, IL-4 and TNFα.

Discussion

The paucity of effective strategies for predicting and preventing PTB has led to concerted research efforts in recent years to identify and characterize biomarkers associated with preterm delivery. In parallel, the vaginal microbiome field has revealed that changes in the vaginal microbiota can precede the onset of PTB [14, 18, 44, 45]. Most of the studies assessing the association between microbiome composition and pregnancy complications such as PTB have focused on women with African or European ancestries, and intriguingly revealed that different microbial taxa can be associated with PTB in each population [14, 18, 44, 45]. Hence, there is a need to conduct ethnicity-specific studies in order to understand whether there is a universal predictive signature, or rather whether different variations in the vaginal microbiome are linked with risk of PTB in distinct populations. Prior to this study, there was no knowledge linking vaginal microbiome or immune parameters with PTB in an entirely Asian cohort or in any Burman or Karen populations.

The women enrolled in this study had no prior adverse obstetric or overt medical history, and therefore represented the women in whom a predictive signature would be of highest value in identifying those apparently low-risk individuals at increased risk of unanticipated PTB. Our data revealed that this cohort of Asian women carried a different vaginal microbial composition compared to women of European and American ancestries [46, 47]. Although the cause of these differences remains unclear, both genetic and environmental factors, including geographic location, diet, age, BMI, drug exposure, physical activities, and availability of resources such as access to medical care are likely to contribute [48-51]. Thus our results confirm the need for ethnicity-specific studies to identify microbial signatures associated with healthy and complicated pregnancies, including the risk of PTB.

While a normal Nugent score is always thought to be accompanied by a healthy pregnancy outcome, our data show that 60% of Asian women in the TB group had an intermediate Nugent score throughout their pregnancy: thus in these women, an intermediate Nugent score was not associated with adverse outcome. The high prevalence of intermediate scoring might be due to the higher counts of *L. iners* [52] in their vaginal microbiome, which has been previously reported as the most dominant *Lactobacillus spp.* detected in Asian women [18]. The presence *L. iners* in the vaginal microbiome has been also strongly associated with an intermediate transient microbiota characterized by an intermediate Nugent score of 4–6 [53]. A previous

study on Caucasian women reported that women dominated by *L. iners* during the first trimester of pregnancy were 10 times more likely than those carrying other *Lactobacilli species* to transition to a dysbiotic microbiome during pregnancy and to deliver preterm [54], but we do not see this here. So while there is still a lot of controversy over whether *L. iners* is more likely a friend or a foe overall, our study reports for the first time that *L. iners* is associated with healthy pregnancy outcomes in an Asian cohort.

Although there are several studies that have assessed PTB in other ethnicities[13, 14, 18, 54], Asian populations in general have been relatively under-studied, and this is the first study to be conducted on the composition of the vaginal microbiome in Karen and Burman pregnant women and its relation to delivery at full term or preterm. Our data show that in TB women there was a transition in the vaginal microbiome between two major CSTs: CST-1 (*L. crispatus*) and CST-III (*L. iners*) with a minimal representation of other CSTs throughout pregnancy. In contrast, in PTB women, we found a decrease in the abundance of *L. crispatus* and *Finegoldia* coinciding with an increase in the prevalence of CST-IVB taxa, particularly *P. buccalis*. This differential microbial signature was detectable as early as the first trimester and was positively correlated with the high Nugent score observed in the PTB group. *Finegoldia* is usually considered a member of CST-IVA when present in combination with a modest proportion of *Lactobacillus species*, and a low proportion of *Anaerococcus*, *Corynebacterium* or *Streptococcus* [43] [55]: perhaps due to the relatively lower abundance of the remaining members, CST-IVA was minimally detected in our cohort and *Finegoldia* was mostly represented by itself. It is also worth mentioning that CST-IVA is often associated with a low Nugent score [43].

All women enrolled on the study were considered low risk for pregnancy complications including pre-term birth. Accordingly, the majority of deliveries in both PTB and TB cohorts occurred vaginally after spontaneous onset of labor or rupture of membranes, which indicates that the local vaginal environment or production of proinflammatory cytokines in the lead up to labor could play a role in modulation of local vaginal environment and thereby risk of PTB. Given that one of the main aims of this study was to generate a predictive PTB signature, we measured the levels of local cytokines in both groups throughout pregnancy. Our data showed that levels of both IL-4 and IFN-γ were lower in the PTB group compared to women who delivered full term, and this was evident from the first trimester of pregnancy. We also observed that vaginal IL-6 and IL-8 levels were significantly increased during trimester-3 or before

delivery in the PTB group. Increased expression of proinflammatory cytokines, including IL-8, IL-6 and TNF-α has been previously reported [14] and is predicted to play a role in the induction of labor.

Our correlation analysis at the earliest gestational time also revealed two microbial scenarios according to pregnancy outcome: in women who delivered full term, *L. crispatus*, *L. gasseri* and *Finegoldia* were negatively correlated with *L. iners* and with the dysbiotic taxa from the CST-IVB; *L. iners* abundance was also negatively correlated with the dysbiotic taxa. Women who delivered preterm also exhibited a negative correlation between *L. crispatus* and *L. gasseri* with the dysbiotic taxa and with *L. iners*; but here they were also negatively correlated with *Finegoldia*. This was further confirmed by the fact that the relative abundance of *Finegolida* in the PTB group was significantly lower compared to the TB group throughout pregnancy. *Finegolida* was positively loosely correlated with IL-2 and IL-4, both higher in women who delivered full term.

Conclusions

In this paper we show that higher abundance of *P. buccalis* accompanied by lower levels of *L. crispatus* and *Finegoldia* in Asian pregnant women may represent a predictive PTB signature that could be used as early as the first trimester to identify those most at risk. More broadly, our results highlight the importance of ethnicity-specific microbiome studies to assess the risk of PTB, especially in otherwise low-risk women. Our findings have several strengths: first, they allow the prospective identification of apparently low-risk women who actually carry a high risk of PTB - in low-resource settings early identification of PTB risk has a real chance of improving the outcome for mother and child by ensuring that they receive extra monitoring and deliver in a medical setting. Secondly, these data are more broadly applicable to Asian women in ethnically-diverse populations in high income nations, such as Qatar and Singapore, where Burman and Karen women might well be a minority group and specific knowledge of their microbiome could be useful; and thirdly, this study represents an important step towards understanding microbiome-induced complications and exploring the potential of developing personalized anti-microbial therapies targeting specific bacteria and aiming to reduce the risks of PTB. Previous studies have included a minority of Asian women, but this unique signature has so far been “lost in the noise”. However, the question remains whether the vaginal microbiota is the direct cause of high PTB risk, or if there is something else that drives the dysbiosis as well as the PTB? Could it be that immune dysregulation, leading to the observed difference in vaginal cytokine levels is actually the reason both vaginal microbial dysbiosis and subsequently high PTB risk? These could be interesting avenues for investigation for future research.

Declarations

Ethics approval and consent to participate

The study was approved by the Institutional Review Board (IRB) of Sidra Medicine under (IRB protocol #1705010909), by the ethics committee of the faculty of Tropical Medicine, Mahidol University, Thailand (TMEC 15-062), the University of Oxford Central University Research, UK (OxTREC: 33-15). Written informed consent was obtained from all participants before their enrolment in the study. All experiments were performed in accordance with the approved guidelines.

Consent for publication

All authors have reviewed the final version of the manuscript and approved it for publication. The contents of the paper have not been published previously.

Availability of data and material

All data generated in the study are included in this article [and its supplementary information files].

Competing interests

The authors declare that they have no competing interests.

Funding

This project is financially supported by funds from Sidra Medicine to Dr. Souhaila Al Khodor.

Authors' contributions

SK conceived and designed the study. SK, DC, BK, AM, TK, TB, RM and FN designed the cohort. MK, SM, PS, MS and DE performed the experiments. TB, RM and FN recruited and consented the study participants. MK and SK wrote the manuscript with input from co-authors. The authors read and approved the final manuscript.

Acknowledgements

We gratefully acknowledge the team at Shoklo Malaria Research Unit (SMRU), Mae Sot, Thailand and the study participants. We thank Dr. Lucy Robinson of Insight Editing London for editing the manuscript.

References

1. Blencowe, H., et al., *National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications.* Lancet, 2012. **379**(9832): p. 2162-72.
2. Vogel, J.P., et al., *The global epidemiology of preterm birth.* Best Pract Res Clin Obstet Gynaecol, 2018. **52**: p. 3-12.
3. Liu, L., et al., *Global, regional, and national causes of under-5 mortality in 2000-15: an updated systematic analysis with implications for the Sustainable Development Goals.* Lancet, 2016. **388**(10063): p. 3027-3035.
4. Mwaniki, M.K., et al., *Long-term neurodevelopmental outcomes after intrauterine and neonatal insults: a systematic review.* Lancet, 2012. **379**(9814): p. 445-52.
5. Shah, R., et al., *Incidence and risk factors of preterm birth in a rural Bangladeshi cohort.* BMC Pediatr, 2014. **14**: p. 112.
6. Beck, S., et al., *The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity.* Bull World Health Organ, 2010. **88**(1): p. 31-8.
7. Soma-Pillay, P., et al., *Physiological changes in pregnancy.* Cardiovasc J Afr, 2016. **27**(2): p. 89-94.
8. Workalemahu, T., et al., *Genetic and Environmental Influences on Fetal Growth Vary during Sensitive Periods in Pregnancy.* Sci Rep, 2018. **8**(1): p. 7274.
9. Lunde, A., et al., *Genetic and environmental influences on birth weight, birth length, head circumference, and gestational age by use of population-based parent-offspring data.* Am J Epidemiol, 2007. **165**(7): p. 734-41.
10. Lockwood, C.J., *Predicting premature delivery--no easy task.* N Engl J Med, 2002. **346**(4): p. 282-4.
11. Lamont, R.F., *Infection in the prediction and antibiotics in the prevention of spontaneous preterm labour and preterm birth.* BJOG, 2003. **110 Suppl 20**: p. 71-5.
12. Nuriel-Ohayon, M., H. Neuman, and O. Koren, *Microbial Changes during Pregnancy, Birth, and Infancy.* Front Microbiol, 2016. **7**: p. 1031.
13. Serrano, M.G., et al., *Racioethnic diversity in the dynamics of the vaginal microbiome during pregnancy.* Nat Med, 2019. **25**(6): p. 1001-1011.
14. Fettweis, J.M., et al., *The vaginal microbiome and preterm birth.* Nat Med, 2019. **25**(6): p. 1012-1021.
15. Brown, R.G., et al., *Vaginal dysbiosis increases risk of preterm fetal membrane rupture, neonatal sepsis and is exacerbated by erythromycin.* BMC Med, 2018. **16**(1): p. 9.
16. 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, 2017. **114**(37): p. 9966-9971.
17. Petrova, M.I., et al., *Lactobacillus species as biomarkers and agents that can promote various aspects of vaginal health.* Front Physiol, 2015. **6**: p. 81.
18. Ravel, J., et al., *Vaginal microbiome of reproductive-age women.* Proc Natl Acad Sci U S A, 2011. **108 Suppl 1**: p. 4680-7.
19. Ling, Z., et al., *Associations between vaginal pathogenic community and bacterial vaginosis in Chinese reproductive-age women.* PLoS One, 2013. **8**(10): p. e76589.
20. Yoshimura, K., et al., *Intravaginal microbial flora by the 16S rRNA gene sequencing.* Am J Obstet Gynecol, 2011. **205**(3): p. 235 e1-9.
21. Hong, K.H., et al., *Analysis of the Vaginal Microbiome by Next-Generation Sequencing and Evaluation of its Performance as a Clinical Diagnostic Tool in Vaginitis.* Ann Lab Med, 2016. **36**(5): p. 441-9.
22. Brummaier, T., et al., *A prospective cohort for the investigation of alteration in temporal transcriptional and microbiome trajectories preceding preterm birth: a study protocol.* BMJ Open, 2019. **9**(1): p. e023417.

23. Summerer, M., et al., *Large-scale mitochondrial DNA analysis in Southeast Asia reveals evolutionary effects of cultural isolation in the multi-ethnic population of Myanmar.* BMC Evol Biol, 2014. **14**: p. 17.
24. Koullali, B., et al., *The association between parity and spontaneous preterm birth: a population based study.* BMC Pregnancy Childbirth, 2020. **20**(1): p. 233.
25. Eschenbach, D.A., et al., *Diagnosis and clinical manifestations of bacterial vaginosis.* Am J Obstet Gynecol, 1988. **158**(4): p. 819-28.
26. Nugent, R.P., M.A. Krohn, and S.L. Hillier, *Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation.* J Clin Microbiol, 1991. **29**(2): p. 297-301.
27. Mattei, V., et al., *Evaluation of Methods for the Extraction of Microbial DNA From Vaginal Swabs Used for Microbiome Studies.* Front Cell Infect Microbiol, 2019. **9**: p. 197.
28. Andrews, S., *FastQC: a quality control tool for high throughput sequence data.* 2010.
29. Zhang, J., et al., *PEAR: a fast and accurate Illumina Paired-End reAd mergeR.* Bioinformatics, 2014. **30**(5): p. 614-20.
30. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data.* Bioinformatics, 2014. **30**(15): p. 2114-20.
31. Caporaso, J.G., et al., *QIIME allows analysis of high-throughput community sequencing data.* Nature methods, 2010. **7**(5): p. 335-336.
32. McMurdie, P.J. and S. Holmes, *phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data.* PLoS One, 2013. **8**(4): p. e61217.
33. Lozupone, C.A., et al., *Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities.* Appl Environ Microbiol, 2007. **73**(5): p. 1576-85.
34. Sharma, S., et al., *Evaluation of the antimicrobial, antioxidant, and anti-inflammatory activities of hydroxychavicol for its potential use as an oral care agent.* Antimicrob Agents Chemother, 2009. **53**(1): p. 216-22.
35. Witten, D.M., R. Tibshirani, and T. Hastie, *A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis.* Biostatistics, 2009. **10**(3): p. 515-34.
36. Rohart, F., et al., *mixOmics: An R package for 'omics feature selection and multiple data integration.* PLoS Comput Biol, 2017. **13**(11): p. e1005752.
37. Donders, G.G., et al., *Pathogenesis of abnormal vaginal bacterial flora.* Am J Obstet Gynecol, 2000. **182**(4): p. 872-8.
38. Brocklehurst, P., et al., *Antibiotics for treating bacterial vaginosis in pregnancy.* Cochrane Database Syst Rev, 2013(1): p. CD000262.
39. DiGiulio, D.B., et al., *Temporal and spatial variation of the human microbiota during pregnancy.* Proc Natl Acad Sci U S A, 2015. **112**(35): p. 11060-5.
40. Chao, A., *Estimating the population size for capture-recapture data with unequal catchability.* Biometrics, 1987. **43**(4): p. 783-91.
41. Shannon, C.E., *A mathematical theory of communication, Part II.* Bell Syst. Tech. J., 1948. **27**: p. 623-656.
42. Simpson, E.H., *Measurement of Diversity.* Nature, 1949. **163**(4148): p. 688-688.
43. Ma, Z.S. and L. Li, *Quantifying the human vaginal community state types (CSTs) with the species specificity index.* PeerJ, 2017. **5**: p. e3366.
44. Goldenberg, R.L., et al., *Epidemiology and causes of preterm birth.* Lancet, 2008. **371**(9606): p. 75-84.
45. Elovitz, M.A., et al., *Cervicovaginal microbiota and local immune response modulate the risk of spontaneous preterm delivery.* Nat Commun, 2019. **10**(1): p. 1305.
46. MacIntyre, D.A., et al., *The vaginal microbiome during pregnancy and the postpartum period in a European population.* Sci Rep, 2015. **5**: p. 8988.
47. Walther-Antonio, M.R., et al., *Pregnancy's stronghold on the vaginal microbiome.* PLoS One, 2014. **9**(6): p. e98514.

48. Conlon, M.A. and A.R. Bird, *The impact of diet and lifestyle on gut microbiota and human health.* Nutrients, 2014. **7**(1): p. 17-44.
49. Verwijs, M.C., et al., *Impact of oral metronidazole treatment on the vaginal microbiota and correlates of treatment failure.* Am J Obstet Gynecol, 2020. **222**(2): p. 157 e1-157 e13.
50. Kumar, M., et al., *Effect of DS-2969b, a novel GyrB inhibitor, on rat and monkey intestinal microbiota.* Anaerobe, 2018. **51**: p. 120-123.
51. Singh, P., M. Kumar, and S. Al Khodor, *Vitamin D Deficiency in the Gulf Cooperation Council: Exploring the Triad of Genetic Predisposition, the Gut Microbiome and the Immune System.* Front Immunol, 2019. **10**: p. 1042.
52. Jespers, V., et al., *Quantification of bacterial species of the vaginal microbiome in different groups of women, using nucleic acid amplification tests.* BMC Microbiol, 2012. **12**: p. 83.
53. Shipitsyna, E., et al., *Composition of the vaginal microbiota in women of reproductive age--sensitive and specific molecular diagnosis of bacterial vaginosis is possible?* PLoS One, 2013. **8**(4): p. e60670.
54. Petricevic, L., et al., *Characterisation of the vaginal Lactobacillus microbiota associated with preterm delivery.* Sci Rep, 2014. **4**: p. 5136.
55. Gajer, P., et al., *Temporal dynamics of the human vaginal microbiota.* Sci Transl Med, 2012. **4**(132): p. 132ra52.

Table 1. Description of the cohort included in this paper.

	TB (n=36)	PTB (n=18)	*P*-value*
Age at conception in years; Median (IQR)	24 (21-27)	21.5 (20-24.5)	0.265*
Height at conception in cm mean (IQR)	151.6 (149.2-155.4)	154.15 (150.5-155.6)	0.267*
Weight at conception in kilograms mean (IQR)	48 (44.25-55.25)	48 (42.25-48.875)	0.334*
BMI at conception; Median (IQR)	20.86 (19.33-23.39)	20.125 (18.19 - 20.43)	0.150*
Delivery (%) - Vaginal - Caesarean section	 35 (97.2) 1 (2.8)	 18 (100) 0 (0.0)	 1‡
Outcome EGA (days); Median (IQR)	276.5 (269.75-283)	253.5 (242-254.75)	<0.001*
Birth weight in grams (IQR)	3060 (2907.5-3310)	2265 (1980-2440)	<0.001*

*Mann Whitney U Test; ‡ Fisher Exact test
EGA, Evaluation of Gestational Age; IQR, InterQuartile Range
PTB: preterm birth, TB: term birth

Figure legends

Fig. 1 Overview of the study design.

a, Of the 400 Karen and Burman women enrolled in Molecular Signature in Pregnancy (MSP) study, 18 pregnant women experienced PTB (pre-term birth) and 36 women who delivered at term (TB) were selected for this study. Vaginal swabs were taken from these women at Shoklo Malaria Research Unit (SMRU) health care clinics at each trimester of pregnancy: T-1 (8-14 weeks), T-2 (20-24 weeks), T-3 (32-35 weeks) and at delivery. Detailed demographic information was collected at the first visit. Vaginal swab samples were used for Nugent scoring to determine bacterial vaginosis status, for metagenomics identification of bacterial species present, and for proteomic analysis of cytokine levels; all conducted at Sidra Medicine, Qatar. BV: bacterial vaginosis; OTU: operational taxonomic unit; W, weeks of gestation. **b,** Number of samples collected from women in both groups and processed for each analysis.

Fig. 2 Association between Nugent score and PTB.

a, Representative images of Gram-stained vaginal swabs with scoring normal, intermediate, and BV, Slides were evaluated under 1,000x magnification according to the 10-point Nugent scale. **b,** Heatmap showing the Nugent scores of individual women who delivered at term (TB) and women who experienced pre-term birth (PTB). Columns represent the trimester of pregnancy. Each row represents one subject. **c,** Comparison of the average Nugent score in the TB or PTB groups during pregnancy. *P*-value was calculated using unpaired *t*-test (two-tailed) with Welch's correction for difference in Nugent score between TB and PTB groups. **p* <0.05. T-1: Trimester-1, T-2: Trimester-2; T-3: Trimester-3; N: Normal; I: Intermediate; BV: Bacterial Vaginosis.

Fig. 3 Vaginal microbiome composition in women with TB and PTB

a, Stacked bar plots showing the relative abundance (%) of each microbial species in vaginal swabs from women who had full term birth (TB) and women who experienced preterm birth (PTB). Each vertical bar represents one woman. **b,** Alpha diversity of the microbiome at delivery was compared between the two groups by the number of operational taxonomic units (OTUs) observed and by the Chao1, Shannon and Simpson diversity indices. The asterisks indicate a significant difference in diversity of microbial communities between the two groups (****P*< 0.001). **c,** Beta diversity plot showing microbial communities clustered using Principle Coordinates Analysis (PCoA) based on Bray–Curtis dissimilarities between vaginal

microbiomes. Statistical significance of the alpha diversity measures was calculated using the Kruskal-Wallis test for non-parametric data, while analysis of similarities method (ANOSIM) was used for calculation of the distance matrix difference between the TB and PTB groups using unweighted beta diversity parameters. P-values lower than 0.05 were considered statistically significant.

Figure 4. Vaginal microbiota profiles of women who had PTB and TB deliveries.

a, Hierarchical clustering of Euclidean distance matrices with Ward linkage on relative abundances of reads for each OTU within individual vaginal swab samples collected at all time points. **b,** Community state types (CST) identified across all the study subjects. Each CST is represented by a different color according to the key shown underneath. **c,** Gestational age category (PTB shown in red, TB shown in blue). **d,** Heatmap of relative abundances of bacterial species within the vaginal microbiota of each woman. Each column represents a woman's vaginal microbiota profile, and each row represents a bacterial species. Only species that represent at least 0.5% of the total microbiome in at least one sample are shown. **e,** Shannon diversity indices calculated for each sample. Each CST is represented by a different color: CST-I: *Lactobacillus crispatus*-dominated; CST-II: *Lactobacillus gasseri*-dominated; CST-III: *Lactobacillus iners*-dominated; CST-IVA: lower abundance of *Lactobacillus spp* together with low proportions of anaerobic bacteria such as *Anaerococcus, Corynebacterium, and Streptococcus*; CST-IVB: dominated by higher abundance of the genera *Atopobium, Prevotella, Parvimonas, Sneathia, Gardnerella, Mobiluncus,* or *Peptoniphilus* and several other taxa.

Fig. 5 Vaginal community state types during the course of pregnancy

a, Stacked area charts of community state type (CST) showing the dynamics of the vaginal microbiome in the full term birth (TB) and preterm birth (PTB) groups at the three trimesters of pregnancy (T-1, T-2, T-3) and at delivery. X-axis represents the gestational age T-1: Trimester-1; T-2: Trimester-2; T-3: Trimester-3, y-axis represents the percentage of each CST in the samples from each group. **b,** Profiles of community state type (CST) for pregnant women who delivered at term (TB) and those who had preterm birth (PTB) as a function of gestational age. Delivery is indicated by a red cross.

Fig. 6 Different bacterial taxa associated with TB and PTB

a, Differences in relative abundance of the top 11 microbial taxa found between full term birth (TB) and preterm birth (PTB) groups in the first trimester; **b,** second trimester; **c,** third trimester and **d,** at the time of delivery. Blue bars represent the TB group, while orange bars represent PTB. The asterisks indicate a significant difference between two groups (*P< 0.05, **P< 0.01, ****P <0.0001). **e,** Longitudinal trends in relative abundance of the statistically significant microbial taxa identified by comparing TB and PTB groups, analyzed using ggplots. Blue lines represent the TB group and red lines the PTB group. The *P*-values were calculated using the unpaired *t*-test (two-tailed) with Welch's correction for difference in proportional microbial abundance between TB and PTB groups.

Fig. 7 Vaginal cytokine levels during PTB and TB pregnancies.
Cytokines were measured in fluid used to elute vaginal swabs from full term birth (TB) and preterm birth (PTB) women in the three trimesters of pregnancy (T-1, T-2, T-3) and at delivery. Blue bars represent the cytokine levels measured in the TB group, while the orange bars represent the cytokine levels in the PTB group. *P*-values were calculated using the unpaired t-test with Welch's correction. The asterisks indicate a significant difference in cytokine levels between the two groups (*P<0.05, **P<0.01, ***P< 0.001 and ****P <0.0001).

Fig. 8 Canonical correlation analysis of vaginal microbial signature and cytokine levels.
The vaginal microbial taxonomic profiles and cytokine levels in samples collected from women who experienced **a,** TB and **b,** PTB at the first trimester (8-14 weeks) were log-transformed and co-integrated using canonical correlation analysis. Cytokines are represented as red diamonds, and bacteria are represented as blue circles. Positively correlated variables are grouped together, while negatively correlated variables are positioned on opposite sides of the plot origin. Thus cytokines or microbial taxa that are clustered tightly are highly correlated, and factors that are distant from each other are not correlated. Pbuc: *Prevotella buccalis*; Fine: *Finegoldia*; Chla: *Chlamydia trachomatis*; Lcri: *Lactobacillus crispatus*; Line: *Lactobacillus iners*; Lgas: *Lactobacillus gasseri;* Ljen: *Lactobacillus jensenii;* Prev: *Prevotella* 6; *Ssan: Sneathia sanguinegens*; Upar: *Ureaplasma parvum; Sam: Sneathia amnii*

Supplementary Fig. 1. Vaginal microbial diversity analysis based on ethnicity

a, Beta diversity analysis to estimate the dissimilarity and similarity of bacterial communities between Burman and Karen population. Principal coordinates analysis (PCoA) derived from the dissimilarity matrix of unweighted UniFrac distance. Statistical analysis calculated using anosim analysis. **b,** Alpha diversity of the microbiome in all samples collected was measured between the two ethnic groups by the number of operational taxonomic units (OTUs) observed and by the Chao1, Shannon and Simpson diversity indices.

Supplementary Fig. 2. Bacterial Nugent score during pregnancy

Stacked area charts representing the percentage of women with full term birth (TB) and (PTB) with either normal, intermediate or bacterial vaginosis (BV) Nugent scores in each trimester (T-1, T-2, and T-3) and at delivery.

Supplementary Fig. 3. Longitudinal vaginal microbiome profile at phylum level during pregnancy.

Each phylum is shown in a different color. Columns marked on the x-axis represent samples from individual women in the full term birth (TB, upper row) and preterm birth (PTB, lower row) cohorts; individual plots represent samples at each trimester (T-1, T-2, and T-3) and at delivery.

Supplementary Fig. 4. Canonical correlation analysis of vaginal microbial signature and cytokine levels at the time of delivery.

The vaginal microbial taxonomic profiles and cytokine levels in samples collected from women who experienced **a,** TB and **b,** PTB at the delivery were log-transformed and co-integrated using canonical correlation analysis. Pbuc: *Prevotella buccalis*; Fine: *Finegoldia*; Chla: *Chlamydia trachomatis*; Lcri: *Lactobacillus crispatus*; Line: *Lactobacillus iners*; Lgas: *Lactobacillus gasseri;* Ljen: *Lactobacillus jensenii;* Prev: *Prevotella* 6; *Ssan: Sneathia sanguinegens*; Upar: *Ureaplasma parvum; Sam: Sneathia amnii*

Figures

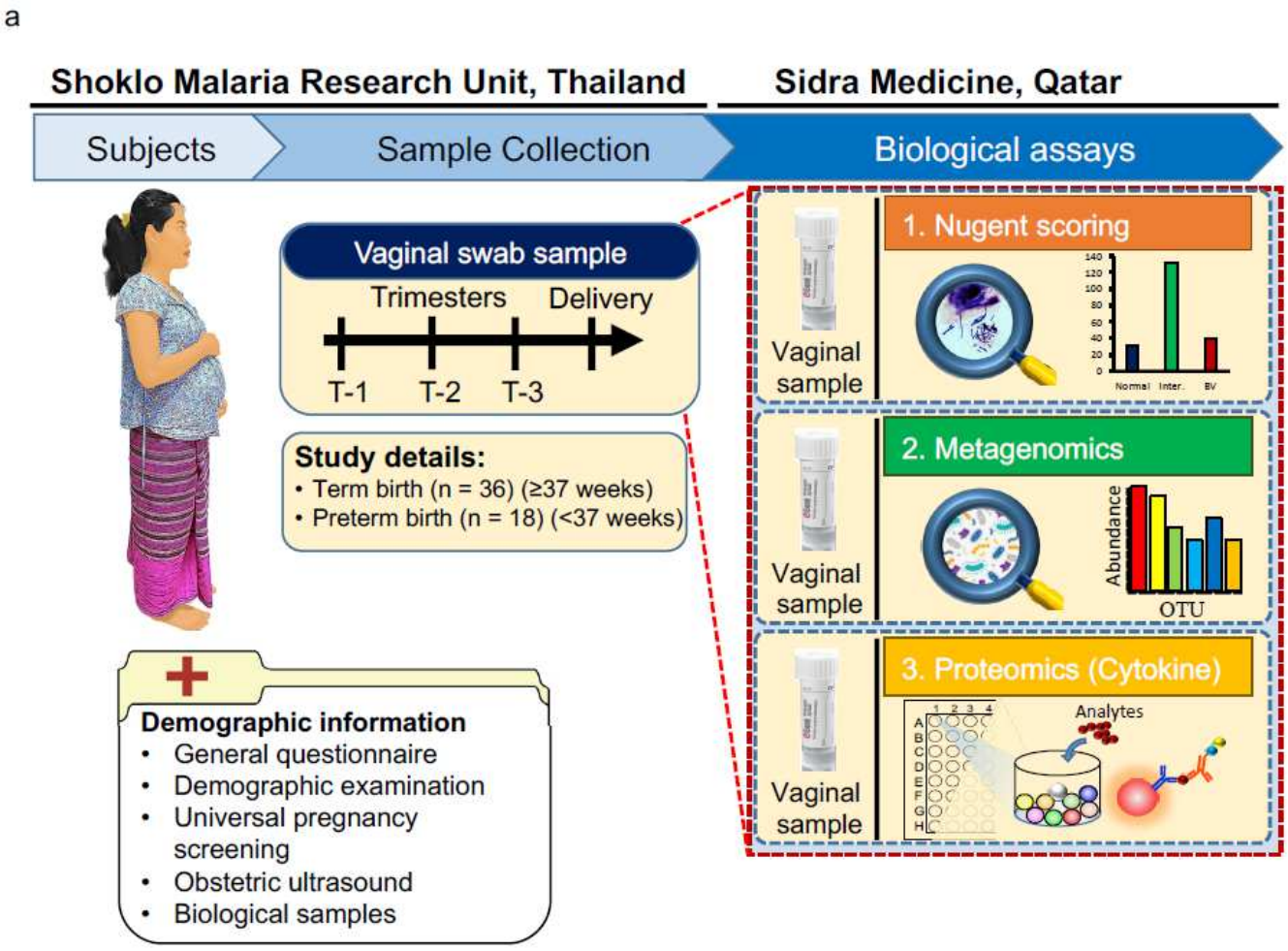

Figure 1

Overview of the study design. a, Of the 400 Karen and Burman women enrolled in Molecular Signature in Pregnancy (MSP) study, 18 pregnant women experienced PTB (pre-term birth) and 36 women who delivered at term (TB) were selected for this study. Vaginal swabs were taken from these women at

Shoklo Malaria Research Unit (SMRU) health care clinics at each trimester of pregnancy: T-1 (8-14 weeks), T-2 (20-24 weeks), T-3 (32-35 weeks) and at delivery. Detailed demographic information was collected at the first visit. Vaginal swab samples were used for Nugent scoring to determine bacterial vaginosis status, for metagenomics identification of bacterial species present, and for proteomic analysis of cytokine levels; all conducted at Sidra Medicine, Qatar. BV: bacterial vaginosis; OTU: operational taxonomic unit; W, weeks of gestation. b, Number of samples collected from women in both groups and processed for each analysis.

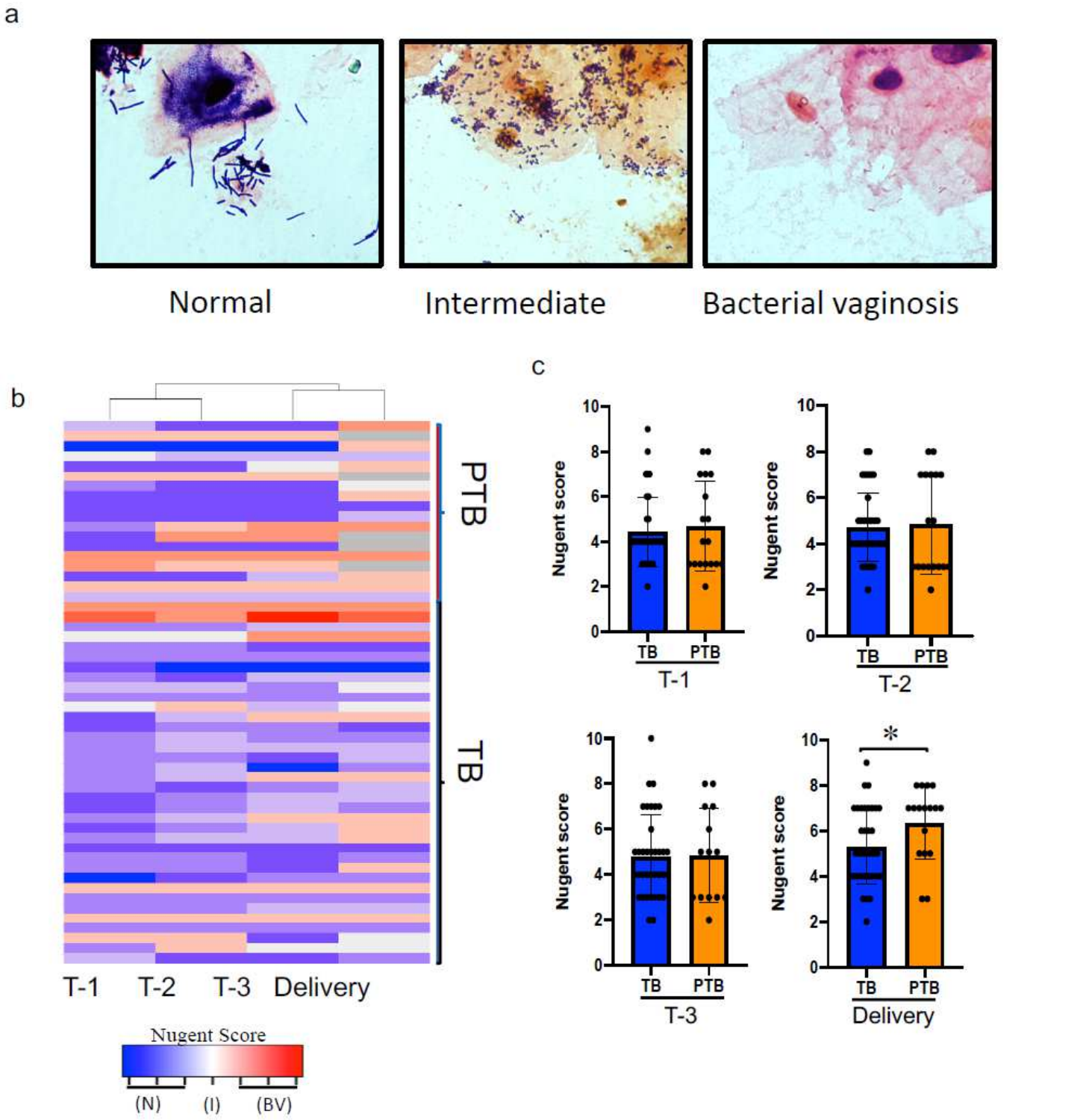

Figure 2

Association between Nugent score and PTB. a, Representative images of Gram-stained vaginal swabs with scoring normal, intermediate, and BV, Slides were evaluated under 1,000x magnification according to the 10-point Nugent scale. b, Heatmap showing the Nugent scores of individual women who delivered at term (TB) and women who experienced pre-term birth (PTB). Columns represent the trimester of pregnancy. Each row represents one subject. c, Comparison of the average Nugent score in the TB or PTB groups during pregnancy. P-value was calculated using unpaired t-test (two-tailed) with Welch's correction for difference in Nugent score between TB and PTB groups. *p <0.05. T-1: Trimester-1, T-2: Trimester-2; T-3: Trimester-3; N: Normal; I: Intermediate; BV: Bacterial Vaginosis.

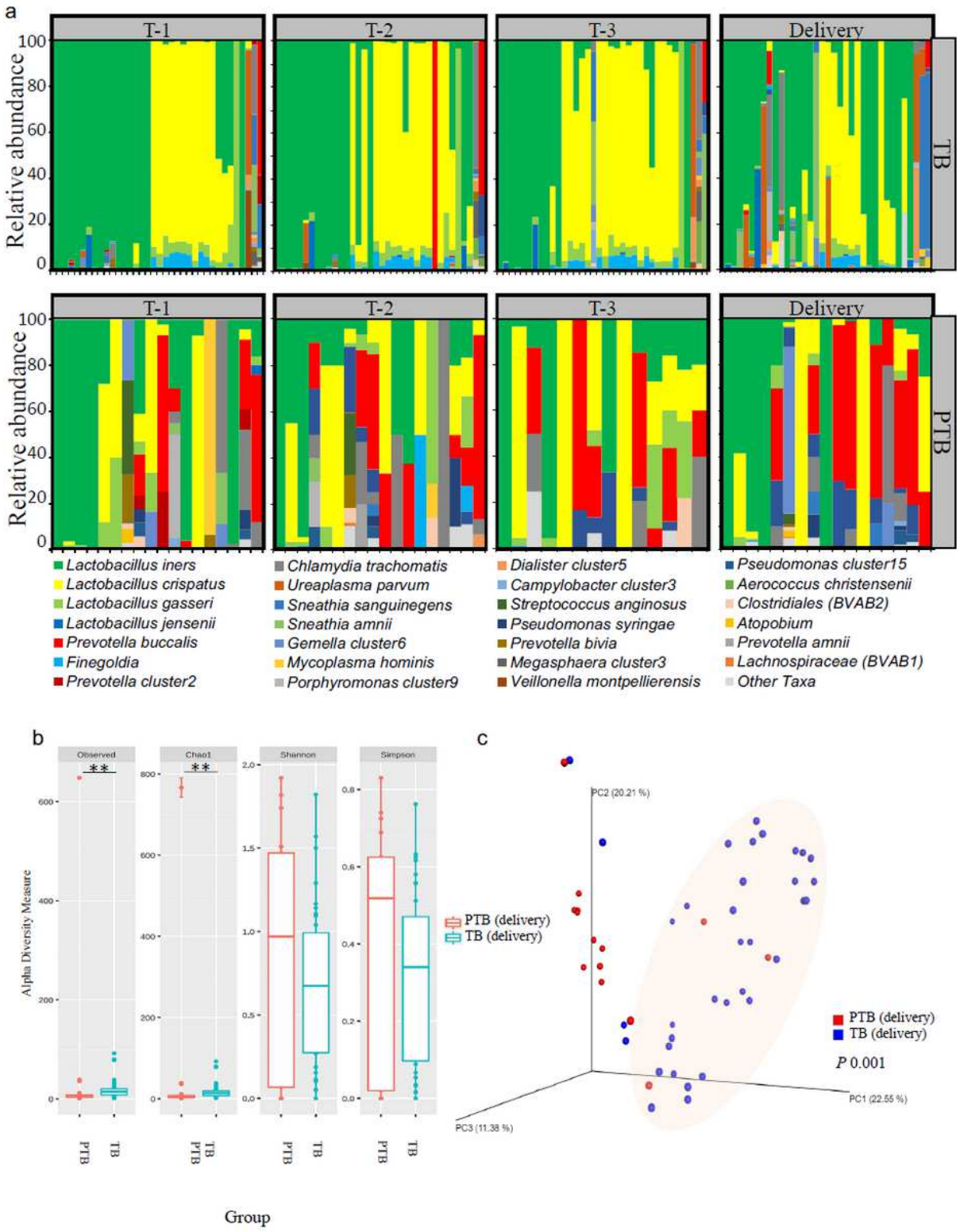

Figure 3

Vaginal microbiome composition in women with TB and PTB a, Stacked bar plots showing the relative abundance (%) of each microbial species in vaginal swabs from women who had full term birth (TB) and women who experienced preterm birth (PTB). Each vertical bar represents one woman. b, Alpha diversity of the microbiome at delivery was compared between the two groups by the number of operational taxonomic units (OTUs) observed and by the Chao1, Shannon and Simpson diversity indices. The

asterisks indicate a significant difference in diversity of microbial communities between the two groups (***P< 0.001). c, Beta diversity plot showing microbial communities clustered using Principle Coordinates Analysis (PCoA) based on Bray–Curtis dissimilarities between vaginal microbiomes. Statistical significance of the alpha diversity measures was calculated using the Kruskal-Wallis test for non-parametric data, while analysis of similarities method (ANOSIM) was used for calculation of the distance matrix difference between the TB and PTB groups using unweighted beta diversity parameters. P-values lower than 0.05 were considered statistically significant.

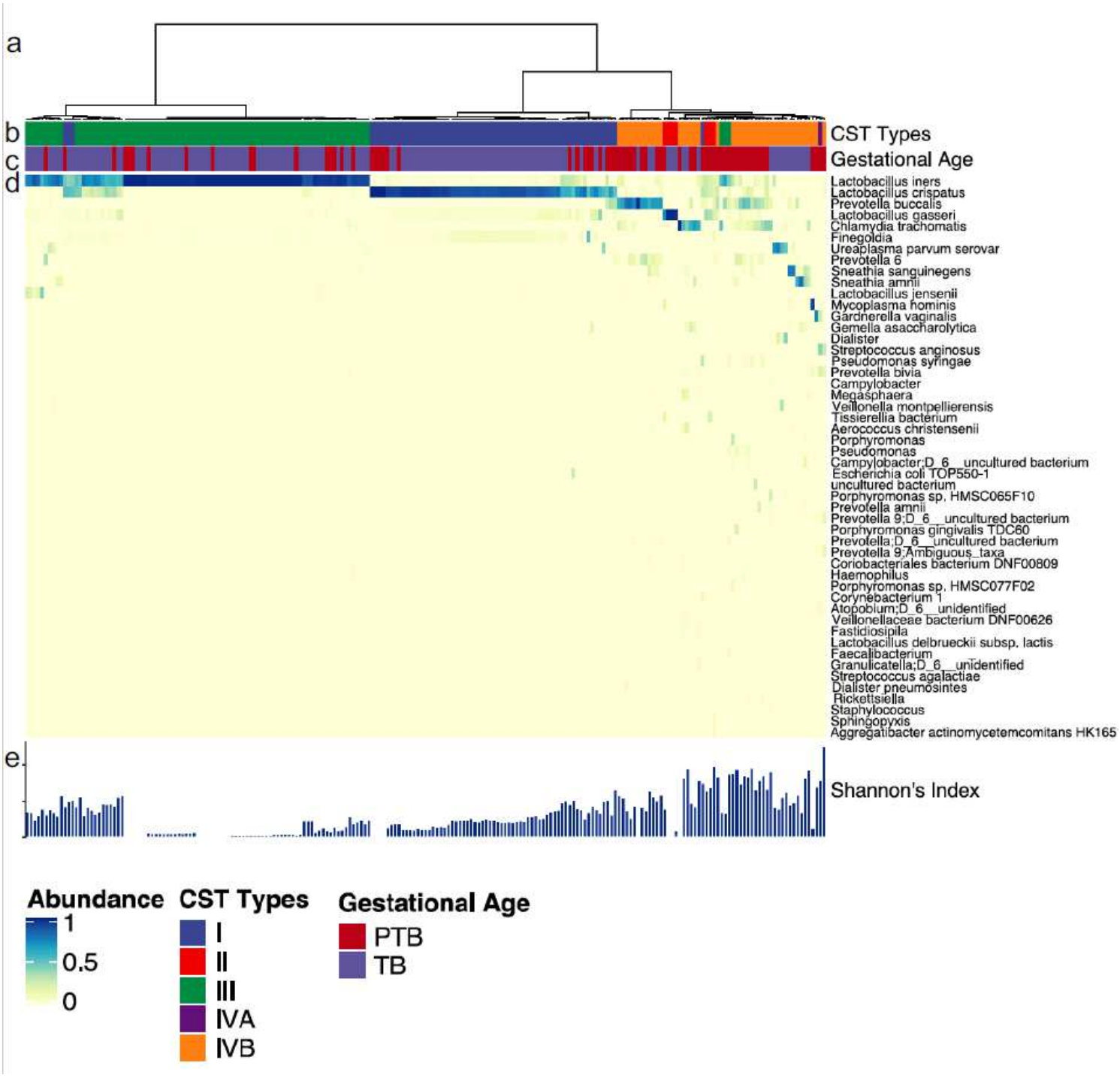

Figure 4

Vaginal microbiota profiles of women who had PTB and TB deliveries. a, Hierarchical clustering of Euclidean distance matrices with Ward linkage on relative abundances of reads for each OTU within individual vaginal swab samples collected at all time points. b, Community state types (CST) identified across all the study subjects. Each CST is represented by a different color according to the key shown underneath. c, Gestational age category (PTB shown in red, TB shown in blue). d, Heatmap of relative abundances of bacterial species within the vaginal microbiota of each woman. Each column represents a woman's vaginal microbiota profile, and each row represents a bacterial species. Only species that represent at least 0.5% of the total microbiome in at least one sample are shown. e, Shannon diversity indices calculated for each sample. Each CST is represented by a different color: CST-I: Lactobacillus crispatus-dominated; CST-II: Lactobacillus gasseri-dominated; CST-III: Lactobacillus iners-dominated; CST-IVA: lower abundance of Lactobacillus spp together with low proportions of anaerobic bacteria such as Anaerococcus, Corynebacterium, and Streptococcus; CST-IVB: dominated by higher abundance of the genera Atopobium, Prevotella, Parvimonas, Sneathia, Gardnerella, Mobiluncus, or Peptoniphilus and several other taxa.

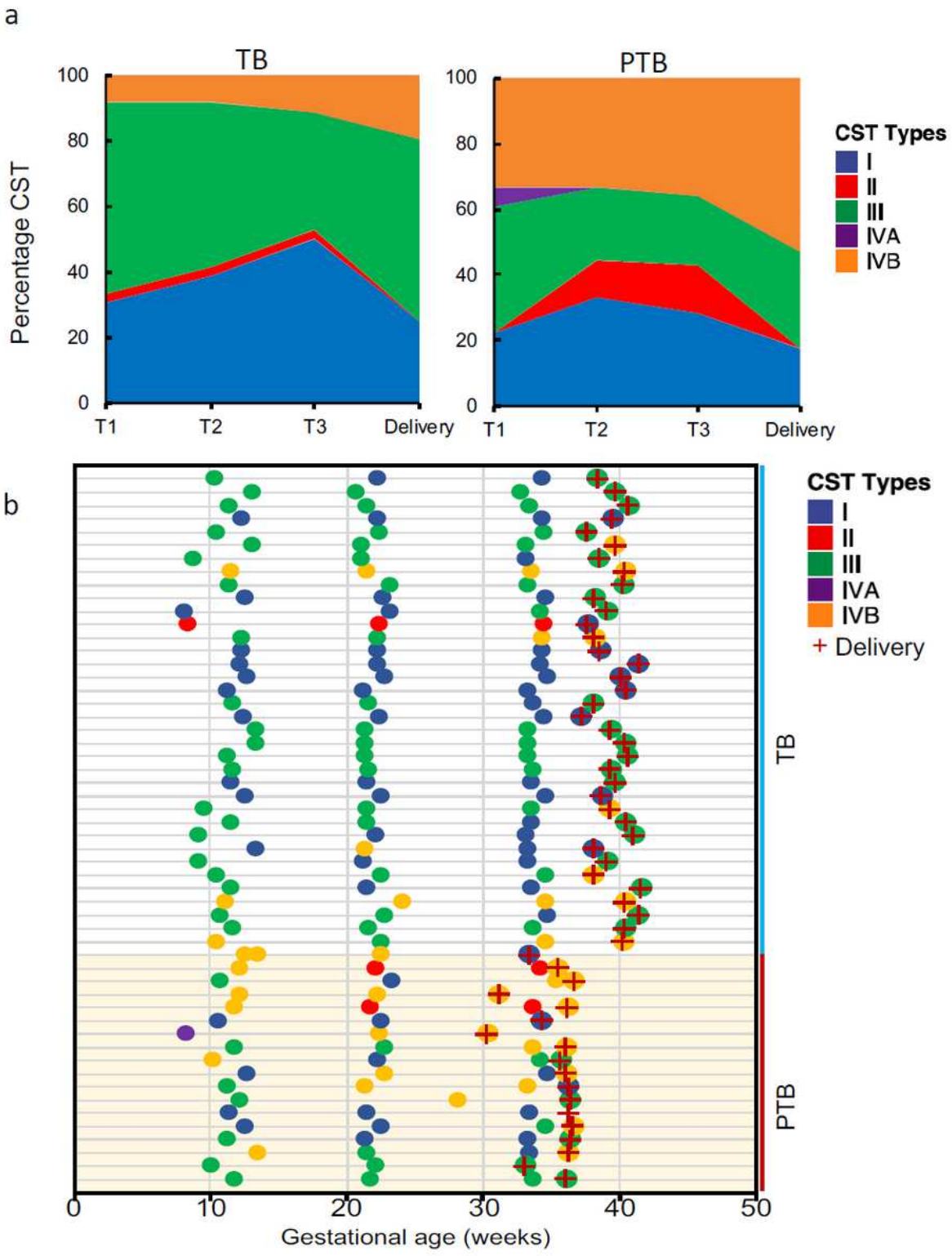

Figure 5

Vaginal community state types during the course of pregnancy a, Stacked area charts of community state type (CST) showing the dynamics of the vaginal microbiome in the full term birth (TB) and preterm birth (PTB) groups at the three trimesters of pregnancy (T-1, T-2, T-3) and at delivery. X-axis represents the gestational age T-1: Trimester-1; T-2: Trimester-2; T-3: Trimester-3, y-axis represents the percentage of each CST in the samples from each group. b, Profiles of community state type (CST) for pregnant women who

delivered at term (TB) and those who had preterm birth (PTB) as a function of gestational age. Delivery is indicated by a red cross.

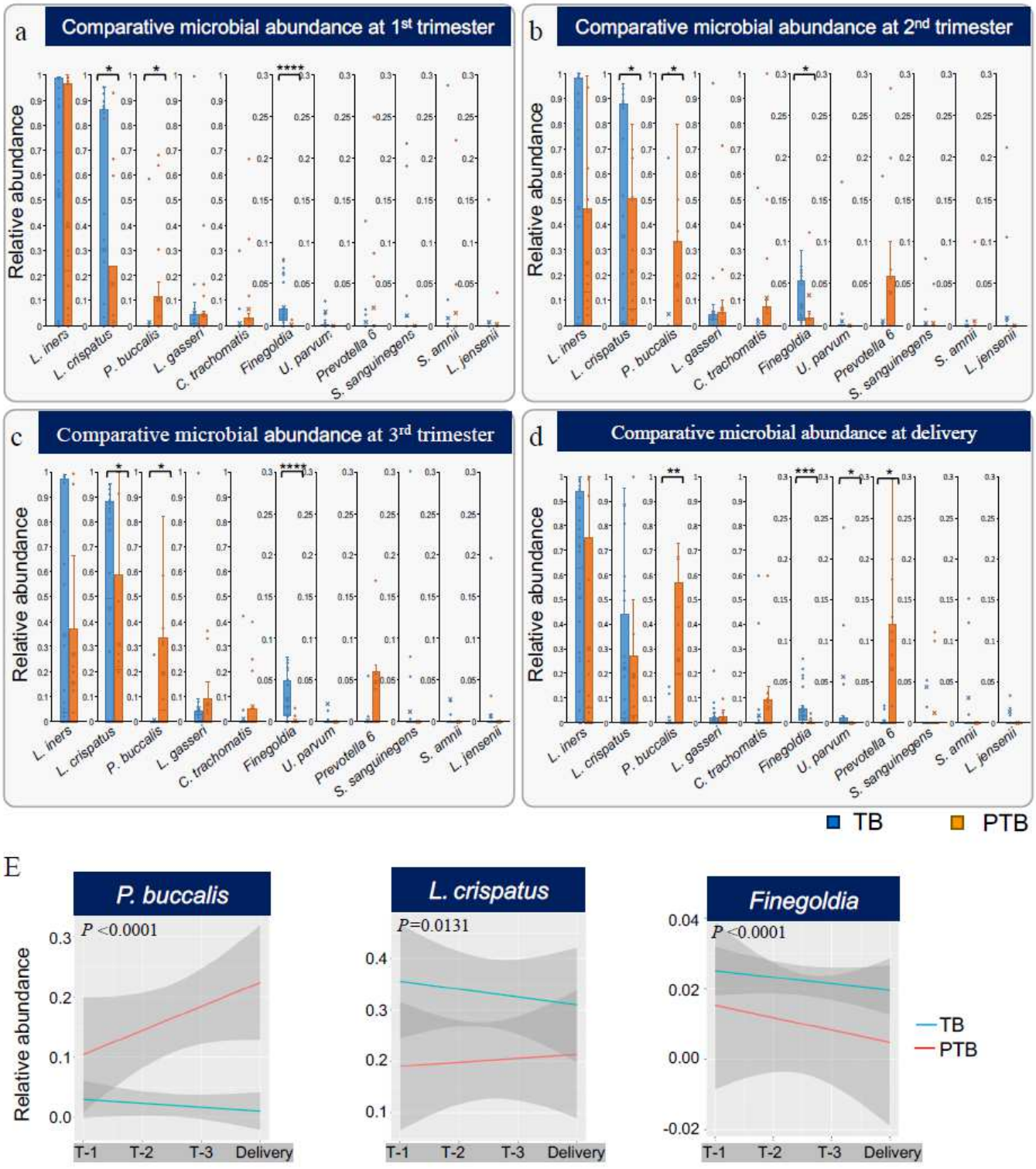

Figure 6

Different bacterial taxa associated with TB and PTB a, Differences in relative abundance of the top 11 microbial taxa found between full term birth (TB) and preterm birth (PTB) groups in the first trimester; b, second trimester; c, third trimester and d, at the time of delivery. Blue bars represent the TB group, while

orange bars represent PTB. The asterisks indicate a significant difference between two groups (*P< 0.05, **P< 0.01, ****P <0.0001). e, Longitudinal trends in relative abundance of the statistically significant microbial taxa identified by comparing TB and PTB groups, analyzed using ggplots. Blue lines represent the TB group and red lines the PTB group. The P-values were calculated using the unpaired t-test (two-tailed) with Welch's correction for difference in proportional microbial abundance between TB and PTB groups.

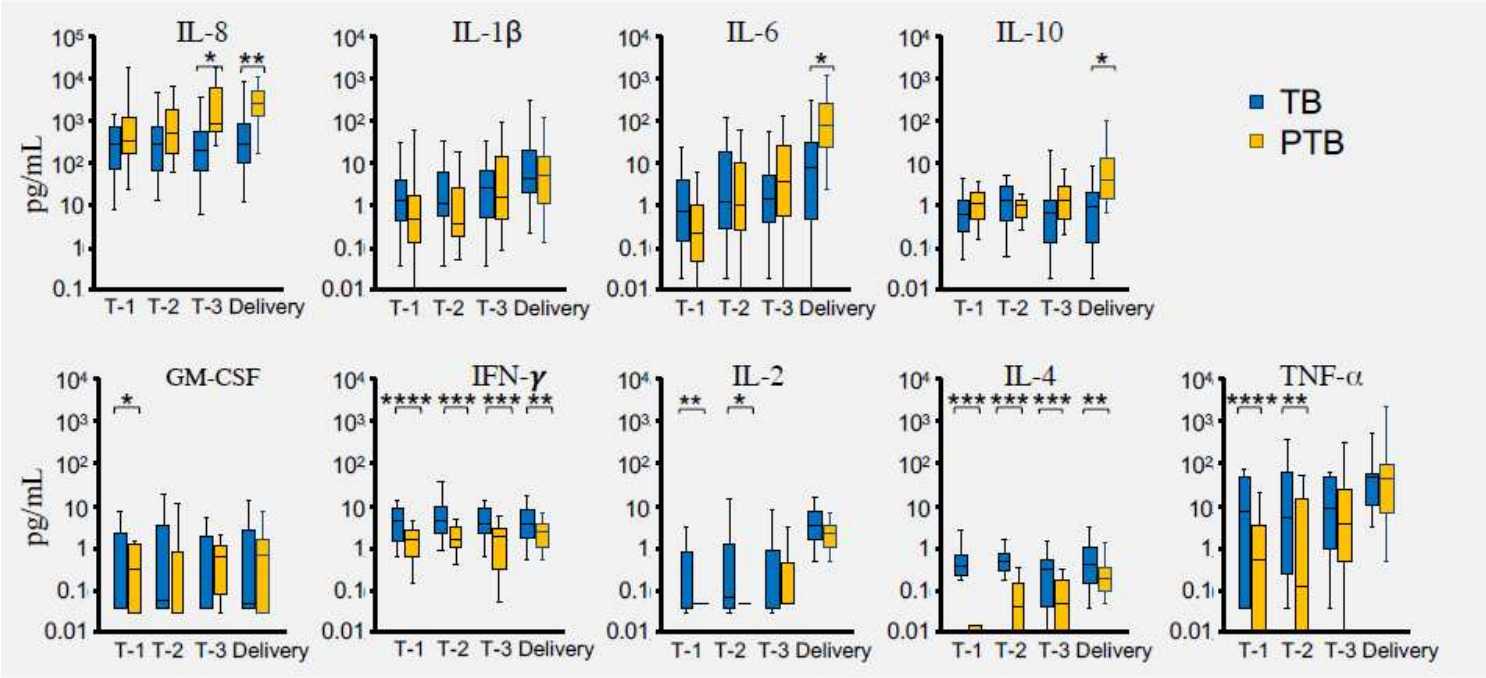

Figure 7

Vaginal cytokine levels during PTB and TB pregnancies. Cytokines were measured in fluid used to elute vaginal swabs from full term birth (TB) and preterm birth (PTB) women in the three trimesters of pregnancy (T-1, T-2, T-3) and at delivery. Blue bars represent the cytokine levels measured in the TB group, while the orange bars represent the cytokine levels in the PTB group. P-values were calculated using the unpaired ttest with Welch's correction. The asterisks indicate a significant difference in cytokine levels between the two groups (*P<0.05, **P<0.01, ***P< 0.001 and ****P <0.0001).

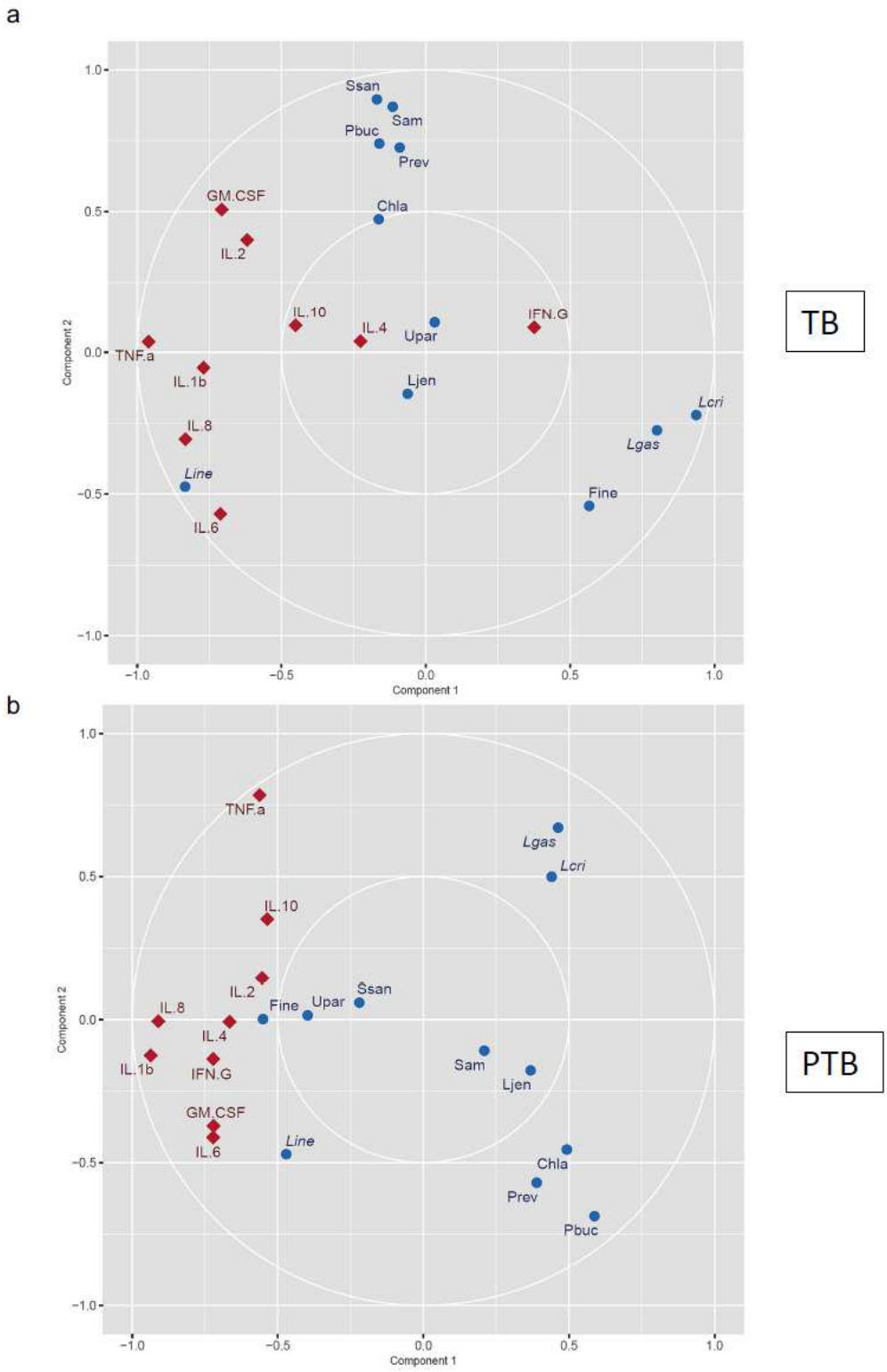

Figure 8

Canonical correlation analysis of vaginal microbial signature and cytokine levels. The vaginal microbial taxonomic profiles and cytokine levels in samples collected from women who experienced a, TB and b, PTB at the first trimester (8-14 weeks) were log-transformed and co-integrated using canonical correlation analysis. Cytokines are represented as red diamonds, and bacteria are represented as blue circles. Positively correlated variables are grouped together, while negatively correlated variables are positioned

on opposite sides of the plot origin. Thus cytokines or microbial taxa that are clustered tightly are highly correlated, and factors that are distant from each other are not correlated. Pbuc: Prevotella buccalis; Fine: Finegoldia; Chla: Chlamydia trachomatis; Lcri: Lactobacillus crispatus; Line: Lactobacillus iners; Lgas: Lactobacillus gasseri; Ljen: Lactobacillus jensenii; Prev: Prevotella 6; Ssan: Sneathia sanguinegens; Upar: Ureaplasma parvum; Sam: Sneathia amnii

Supplementary Files

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

- Table1.pdf
- TableS1.pdf
- Graphicalabstract.pdf
- SIFigures.pdf