

ARTICLE OPEN



Large-scale characterisation of the pregnancy vaginal microbiome and sialidase activity in a low-risk Chinese population

Sherriane Ng^{1,2,8}, Muxuan Chen^{3,4,8}, Samit Kundu^{1,2}, Xuefei Wang⁵, Zuyi Zhou^{3,4}, Zhongdaixi Zheng⁶, Wei Qing^{3,4}, Huafang Sheng^{3,4}, Yan Wang⁵, Yan He^{3,4}, Phillip R. Bennett^{1,2}, David A. MacIntyre^{1,2}✉ and Hongwei Zhou^{3,4,7}✉

Vaginal microbiota-host interactions are linked to preterm birth (PTB), which continues to be the primary cause of global childhood mortality. Due to population size, the majority of PTB occurs in Asia, yet there have been few studies of the pregnancy vaginal microbiota in Asian populations. Here, we characterized the vaginal microbiome of 2689 pregnant Chinese women using metataxonomics and in a subset ($n = 819$), the relationship between vaginal microbiota composition, sialidase activity and leukocyte presence and pregnancy outcomes. Vaginal microbiota were most frequently dominated by *Lactobacillus crispatus* or *L. iners*, with the latter associated with vaginal leukocyte presence. Women with high sialidase activity were enriched for bacterial vaginosis-associated genera including *Gardnerella*, *Atopobium* and *Prevotella*. Vaginal microbiota composition, high sialidase activity and/or leukocyte presence was not associated with PTB risk suggesting underlying differences in the vaginal microbiota and/or host immune responses of Chinese women, possibly accounting for low PTB rates in this population.

npj Biofilms and Microbiomes (2021)7:89; <https://doi.org/10.1038/s41522-021-00261-0>

INTRODUCTION

There is substantial evidence implicating the pregnancy vaginal microbiota in shaping maternal and neonatal health outcomes^{1–3}. The dominance of the vaginal niche by commensal *Lactobacillus* species is often considered “optimal” due to their ability to prevent pathogen colonization through competitive exclusion, in part achieved through the production of antimicrobial compounds and production of lactic acid^{4,5}. Recent studies have highlighted *L. crispatus* dominance as being protective against preterm birth (PTB)^{2,6–9} and neonatal sepsis following preterm prelabour rupture of membranes¹. By contrast, colonization by *L. iners*^{7,10–12} or *Lactobacillus* species depleted, high diversity compositions are associated with an increased risk of PTB^{2,7,9,10,13–15}. Despite the majority of PTBs (60%) occurring in Asia¹⁶, molecular-based characterization of the vaginal microbiota in pregnancy and its relationship with PTB has largely been restricted to Northern American and European populations. Moreover, ethnicity is now recognized as a potential confounder of the relationship between the vaginal microbiome and PTB, particularly between Caucasian and women of African-descent in North American or European populations^{2,9,10,17}.

The Amsel criteria¹⁸ and Nugent scoring system¹⁹ are commonly used to diagnose Bacterial Vaginosis (BV), a condition characterized by a loss of vaginal lactobacilli and overgrowth of anaerobes, which is associated with a two-fold increased risk of PTB^{20–23}. The Amsel criteria requires microscopy and is subject to potential interobserver bias²⁴, whereas Nugent scoring is limited by the requirement for laboratory access. Enzymatic-based assays for rapid BV diagnosis may offer an objective, point-of-care alternative to BV diagnosis in clinical settings, including during pregnancy^{25,26}. These assays often work by measuring microbial sialidase (neuraminidase), produced by BV-

associated bacteria such as *Gardnerella vaginalis*, which removes sialic acid from sialoglycoconjugates including those on the surface of vaginal epithelial cells, providing a nutrient source and exposing glycan-binding sites for bacterial adhesion. Additionally, sialidase is thought to mediate biofilm formation and the establishment of sub-optimal vaginal microbiota compositions^{27–29}. High vaginal sialidase levels have previously been associated with an increased risk of PTB³⁰ and with failure of cervical cerclage³¹, a procedure used to reinforce the cervical opening in women at risk of preterm delivery due to cervical shortening. Sialidase-producing taxa associated with BV have also been implicated in chorioamnionitis, a risk factor for PTB that is characterized by inflammation of the fetal membranes³². Although BV is not classified as an inflammatory syndrome, the disease has been associated with the presence of vaginal leukocytes, which have been purported to offer predictive value in identifying upper reproductive tract infections^{33–35} and PTB^{36–38}. Quantification of leukocyte counts using vaginal wet mount microscopy could therefore represent an easily accessible and cost-effective method to determine cervicovaginal inflammation³⁹.

In this study, we characterized the bacterial component of the vaginal microbiome in 2689 Chinese women sampled at mid-pregnancy and in a subset ($n = 819$), explored the relationship between vaginal microbiota composition, sialidase activity, and leukocyte presence with risk of PTB.

RESULTS

Study population

A total of 2796 women, with a median (IQR) maternal age of 29 years (26–32 years), met the study inclusion criteria and were

¹Imperial College Parturition Research Group, Imperial College London, London, United Kingdom. ²March of Dimes European Prematurity Research Centre, Imperial College London, London, United Kingdom. ³Department of Laboratory Medicine, Zhujiang Hospital, Southern Medical University, Guangzhou, Guangdong, China. ⁴Microbiome Medicine Center, Zhujiang Hospital, Southern Medical University, Guangzhou, Guangdong, China. ⁵Department of Obstetrics and Gynecology, Nanfang Hospital, Southern Medical University, Guangzhou, China. ⁶School of Public Health, Southern Medical University, Guangzhou, Guangdong, China. ⁷State key laboratory of Organ Failure Research, Southern Medical University, Guangzhou, Guangdong, China. ⁸These authors contributed equally: Sherriane Ng, Muxuan Chen. ✉email: d.macintyre@imperial.ac.uk; biodegradation@gmail.com

recruited between November 2015 to December 2018. The median (IQR) time of sampling was 16^{+4} weeks $^{+days}$ gestation (range 15^{+6} – 17^{+6}) after exclusion of women with samples collected in the third trimester. Of these, 1397 delivered at the same hospital and maternal and neonatal outcomes data were obtained. The median (IQR) gestation at delivery was 39^{+3} (38^{+5} – 40^{+1}) weeks $^{+days}$. The PTB (<37 weeks) rate in the cohort was 5.44% (76/1397) (Supplementary Data 1). The remaining 1356 women delivered elsewhere and due to data protection, pregnancy outcome data were not available. There was no significant difference in the gestation of sampling between women with and without available outcome data (Mann-Whitney Test, $p > 0.05$; Supplementary Fig. 1).

The pregnancy vaginal microbiota of Chinese women

The 2796 sequenced vaginal samples generated a total of 58,582,840 reads with a mean read count of 20,952 per sample. Of these samples, 2689 passed library size and microbiome classification criteria (Supplementary Fig. 1). Only those samples collected in the first and second trimester were included in subsequent analyses (2646/2689, 98.4%) After removal of kit and reagent contaminants, a total of 82 taxa were detected and included in subsequent analyses. Vaginal microbiota profiles were classified into 19 groups based upon the dominant (>30% relative abundance) taxa observed within each sample². At the genus level, the majority of samples were dominated by *Lactobacillus* (2275/2646, 85.98%), which were predominately *L. crispatus* (1058/2646, 39.98%) or *L. iners* (952/2646, 35.98%) dominated (Fig. 1a). The median relative abundance of *L. crispatus* and *L. iners* in these samples was 96.37% and 96.19% respectively and the alpha diversity as measured by the Shannon Index was similar between the two groups (median, 0.22 and 0.24) (Supplementary Table 1). Conversely, microbiota dominated by BV-associated species including *Gardnerella* spp. (220/2646, 8.31%), *Prevotella* spp. (25/2646, 0.94%) and *Atopobium* spp. (26/2646, 0.98%) had lower median relative abundances (<70%) and higher alpha diversity (median Shannon Index >0.85) compared to samples classified as *L. crispatus* or *L. iners* dominated (median Shannon Index <0.25, Kruskal-Wallis Test, $p < 0.001$) (Fig. 1a and Supplementary Table 1). The same relationships between alpha diversity and dominant vaginal taxa profiles were observed when read count data was normalized using trimmed mean of *M* values normalization (Supplementary Fig. 2). When these dominant taxa profiles were grouped into Vaginal Microbiome Groups (VMG) consistent with previously reported “community state types”⁴⁰ or “vagitypes”², the most prevalent was VMG I (*L. crispatus* dominated, 1058/2646, 39.98%), followed by VMG III (*L. iners* dominated, 952/2646, 35.98%), VMG IV-A (BV-associated taxa, 282/2646, 10.66%), VMG II (*L. gasseri* dominated, 137/2646, 5.18%), VMG V (*L. jensenii* dominated, 96/2646, 3.63%) and IV-B (Non-BV taxa, 31/2646, 1.17%). Groups dominated by other *Lactobacillus* species or *Bifidobacterium* species were classified into VMG VI (32/2646, 1.21%) and VII (58/2646, 2.19%) (Fig. 1b and Supplementary Table 2), respectively. The proportions of VMGs in women with and without available pregnancy outcome data (Supplementary Fig. 3) were similar (Fisher’s Exact Test, $p = 0.24$). Additional analyses also showed that vaginal microbiota compositional characteristics and their relationship with clinical outcomes were not impacted by the type of extraction kits used (see Supplementary Fig. 4).

Relative abundance of *L. crispatus* was negatively correlated with most vaginal taxa including *L. iners* ($r = -0.48$, $p < 0.001$). *L. iners* was also negatively correlated to most other vaginal taxa except *Megasphaera* spp. ($r = 0.06$, $p < 0.05$) and *Aerococcus* spp. ($r = 0.08$, $p < 0.01$), where weak positive correlations were observed. Several BV-associated taxa were positively correlated with each other including *Gardnerella* spp., with *Atopobium vaginae* ($r = 0.45$, $p < 0.001$), *Prevotella* spp. ($r = 0.34$, $p < 0.001$),

Aerococcus spp. ($r = 0.34$, $p < 0.001$) and Other Species ($r = 0.33$, $p < 0.001$); *Atopobium vaginae* with *Prevotella* spp. ($r = 0.36$, $p < 0.001$), *Megasphaera* spp. ($r = 0.29$, $p < 0.01$), and *Aerococcus* spp. ($r = 0.31$, $p < 0.001$); and *Prevotella* spp. with the non-BV taxa *Streptococcus anginosus* ($r = 0.31$, $p < 0.001$) and *Anaerococcus* spp. ($r = 0.42$, $p < 0.001$), as well as Other Species ($r = 0.45$, $p < 0.001$) (Fig. 1c).

Sialidase activity and vaginal microbiota structure in pregnancy

High vaginal sialidase activity was strongly associated with increased bacterial alpha diversity ($n = 36$; median (IQR) Shannon index 0.93 (0.54–1.55)) compared to women with low vaginal sialidase activity ($n = 783$; median (IQR) Shannon index 0.32 (0.16–0.77), Mann-Whitney Test, $p < 0.001$) (Fig. 2a). Women with high sialidase activity had significantly higher prevalence of *Lactobacillus* species depleted microbiota (30/36, 83.33% versus 183/783, 23.37%; Fisher’s Exact Test, $p = 2.403e-13$) and VMG IV-A (22/36, 61.11% vs. 58/783, 7.40%; Fisher’s Exact Test, $p = 0.0005$) (Fig. 2b, Supplementary Fig. 5). VMG I (2/36, 5.56% versus 319/783, 40.74%) and II (2/36, 5.56% versus 48/783, 6.13%) were rarely observed in women with high sialidase activity whereas prevalence of VMG III was similar in women with high and low sialidase activity (10/36, 27.78% vs. 288/783, 36.78%). Samples classified as VMG IV-B (11/783, 1.40%), V (35/783, 4.47%), VI (9/783, 1.15%), and VII (15/783, 1.92%) were only associated with low sialidase activity. Differential abundance analysis using LDA and LefSe of samples with high and low sialidase activity confirmed *L. crispatus* and *L. jensenii* as being enriched in low sialidase activity samples whereas high activity was characterized by enrichment with BV-associated genera including *Gardnerella*, *Atopobium*, *Prevotella*, and *Megasphaera* (Fig. 2c). In these women, the higher total relative abundance of BV-type taxa (*Gardnerella* spp., *Atopobium vaginae*, *Prevotella* spp. and *Megasphaera* spp.) was observed compared to women with low sialidase activity (Fig. 2d and Supplementary Table 3). Using metataxonomics-defined *Lactobacillus* depleted (positive) and *Lactobacillus* dominated (negative) as the reference test, sialidase activity in this cohort was found to have low sensitivity (14.08%), high specificity (99.01%), moderate positive predictive value (PPV, 71.51%), and high negative predictive value (NPV, 86.72%) for *Lactobacillus* spp. depleted high diversity vaginal microbiota compositions (Fig. 2e).

Leukocyte presence and vaginal microbiota composition during pregnancy

Women with leukocyte grade III wet mounts were associated with a small but significantly higher proportion of *Lactobacillus* dominated microbiomes (382/486, 78.60%) compared to women with leukocyte grades I (112/165, 67.88%), II (55/82, 67.07%), and IV (57/86, 66.28%; Fisher’s Exact Test, $p = 0.003$). There were also higher proportions of VMG III in women with leukocyte grades III (180/486, 37.03%) and IV (46/86, 53.49%) compared to women with grades I (50/165, 30.30%) and II (22/82, 26.83%; Fisher’s Exact Test, $p = 0.02$). Consistent with this, relative abundance levels of *L. iners* were highest in women with leukocyte grades III and IV (Kruskal-Wallis Test, $p < 0.01$) (Fig. 3b). Similar results were observed when using LDA and LefSe analysis to identify differential taxa between leukocyte low (grades I-II) and leukocyte high (grades III-IV) wet mounts with *L. iners* having the strongest effect size of the differentially abundant features identified (Fig. 3c).

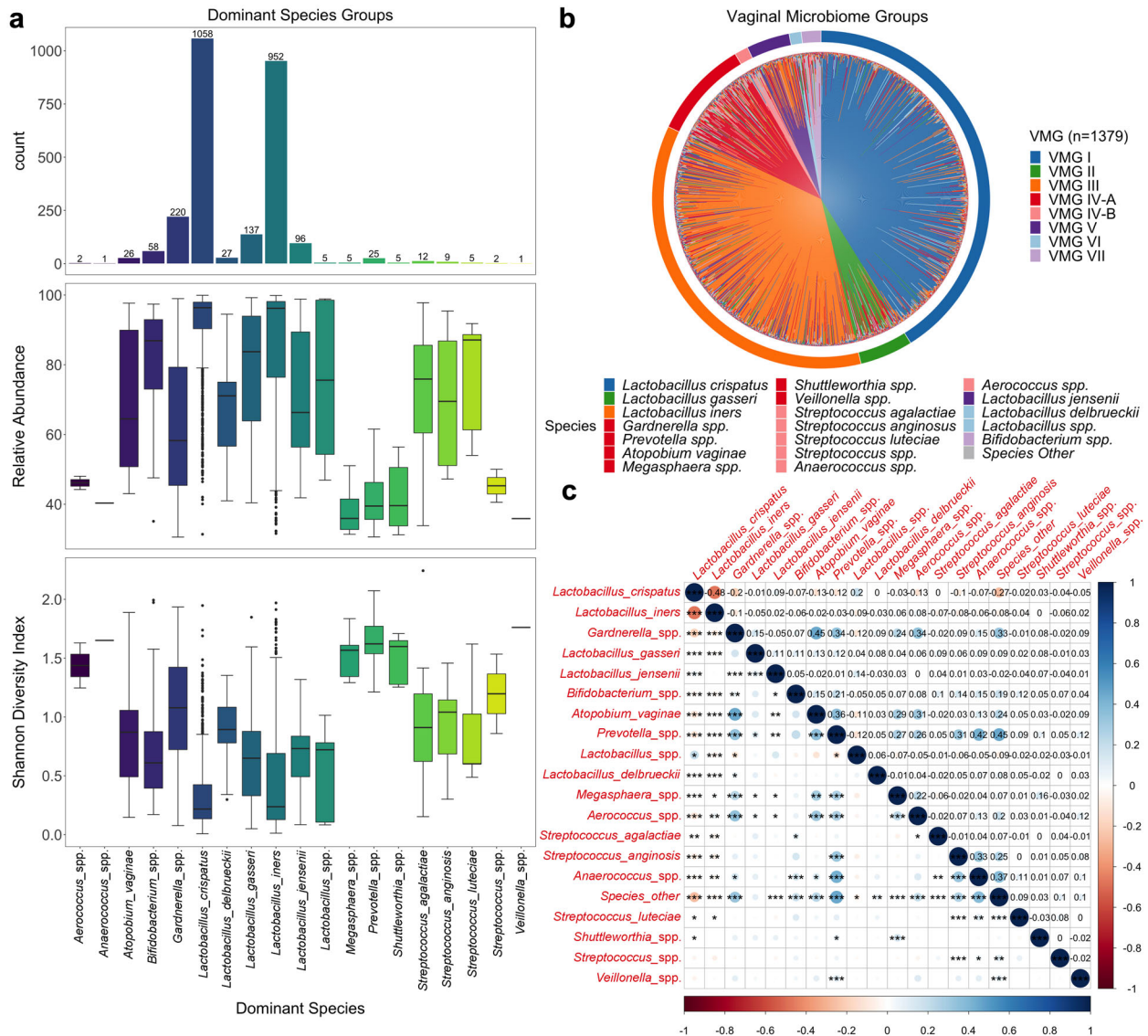


Fig. 1 Vaginal microbiome composition and structure. **a** Bar chart displaying number of samples within each dominant species group. Box plots show relative abundance of species and Shannon diversity index of each dominant species group with bounds of the box representing the first and third quartiles, the center line representing the median and whiskers as min-to-max values. **b** Microbiome profiles of women with outcome data classified by vaginal microbiome groups (VMG). Color of species denotes the VMG group classification with VMG I (*L. crispatus* dominated, blue), II (*L. gasseri* dominated, green), III (*L. iners* dominated, orange), IV-A (bacterial vaginosis associated, red), IV-B (non-BV taxa, pink), V (*L. jensenii* dominated, purple), VI (Other *Lactobacillus*, light blue), and VII (*Bifidobacterium* dominated, light purple). **c** Correlation plot for dominant species showing correlation strength (circles), relationship (positive, blue; negative, red) and significance (***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$).

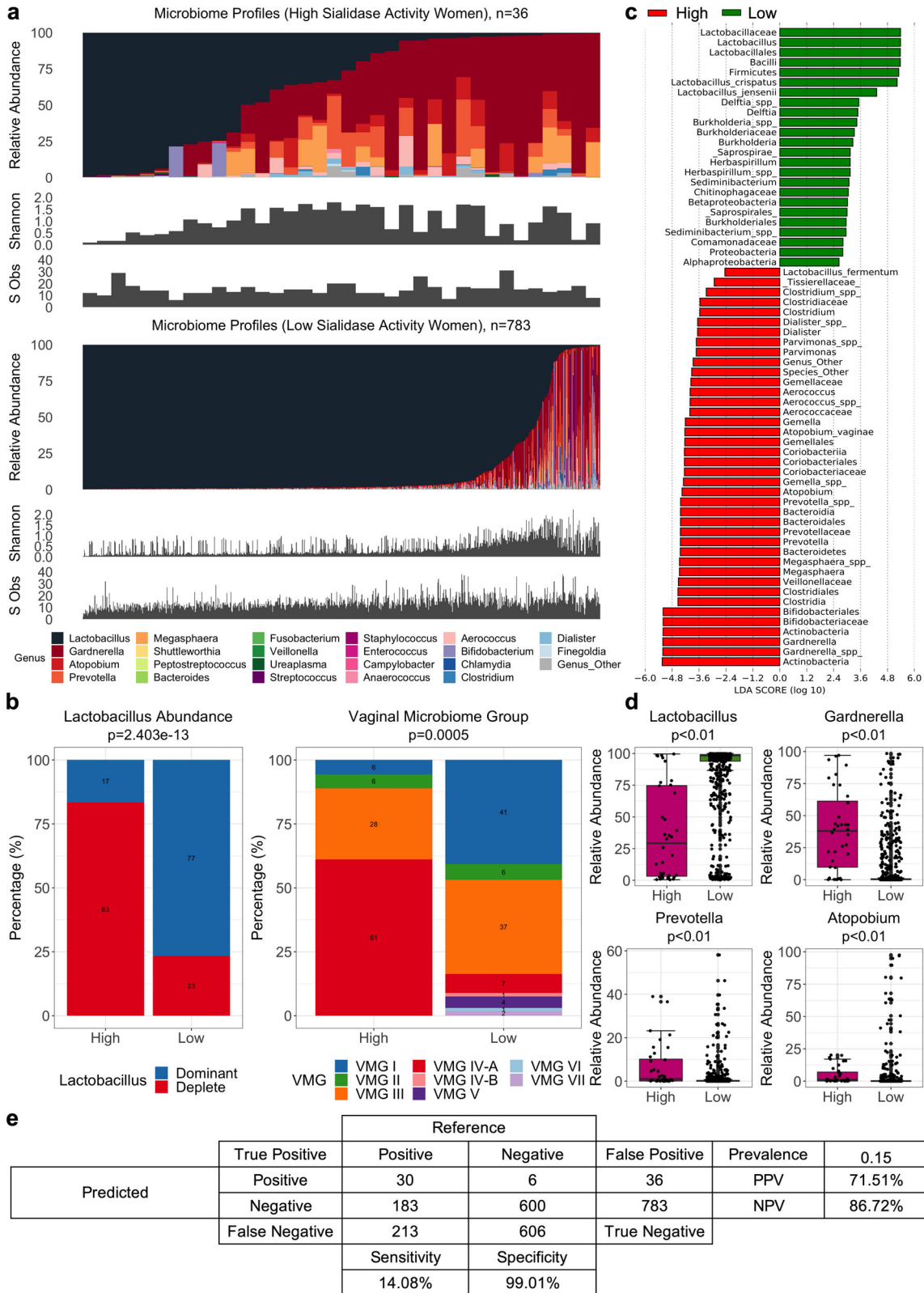
Vaginal microbiota, sialidase activity, and leukocyte presence are not associated with risk of PTB and chorioamnionitis in pregnant Chinese women

In women with available outcome data ($n = 1379$), the proportion of *Lactobacillus* dominated or depleted vaginal microbiota ($p = 0.57$; Fisher's Exact test) and VMGs ($p = 0.05$; Fisher's Exact test) were similar between women delivering preterm or term, including when preterm births were divided into early (<34 weeks gestation) and late (≥ 34 weeks gestation) (Supplementary Fig. 6). Likewise, proportions of high or low sialidase activity ($p = 0.45$; Fisher's Exact test) and leukocyte high/low ($p = 0.74$; Fisher's Exact test) were comparable between women subsequently experiencing preterm or term deliveries (Fig. 4a). Apart from women who had *Lactobacillus* dominated microbiomes, high sialidase activity, and leukocyte low which had insufficient sample size ($n = 2$), no

significant difference was observed in birth gestation based on *Lactobacillus* abundance, VMG prevalence, sialidase activity, and leukocyte wet mount results (Kruskal-Wallis Test, $p > 0.05$) (Fig. 4b and Supplementary Table 4). Logistic regression results similarly showed vaginal microbiome *Lactobacillus* composition ($p = 0.83$), sialidase activity ($p = 0.45$) and leukocyte high or low ($p = 0.70$) did not significantly contribute to birth outcome (Supplementary Table 5). Finally, no relationships were observed between vaginal microbiota composition, sialidase activity and/or leukocyte presence, and chorioamnionitis (Supplementary Fig. 7).

DISCUSSION

In this large cross-sectional study of 2646 pregnant Chinese women, the vaginal microbiota in mid-pregnancy was characterized by *Lactobacillus* spp. dominance and low diversity. This



pattern is consistent with previous descriptions of the healthy mid-pregnancy vaginal microbiome in Caucasian, African and Asian women from European and Northern American populations^{6,10,13,17,41}. Similar results have also been reported in small cohort studies of Chinese and Indian women^{42,43}. One of these

studies examined the vaginal microbiota of 113 Chinese women and reported *L. crispatus* dominance in 45.1% and *L. iners* dominance in 31.9% of women sampled at a mean gestation of 17 weeks⁴². In our cohort, *L. crispatus* (40%) and *L. iners* (36%) dominated vaginal microbiota were also the most frequently

Fig. 2 Comparison between women in sialidase sub-cohort. a Microbiome composition, diversity (Shannon Index) and richness (species observed) for all women with high ($n = 36$) and low ($n = 783$) sialidase activity. **b** Proportion of *Lactobacillus* abundance groups (*Lactobacillus* dominant and *Lactobacillus* deplete) and vaginal microbiome groups (VMG) for women with high or low sialidase activity. Statistical significance based on Fisher's Exact Test. **c** LDA showing effect size of differentially abundant taxa associated with microbiota of women with high (red) or low (green) sialidase activity. **d** Comparison of relative abundance between main BV-associated bacteria identified as significantly different (Mann-Whitney Test, $p < 0.05$) between women with high and low sialidase activity (see Supplementary Table 3). The bounds of the box represent the first and third quartiles, center line represents the median, and whiskers show min-to-max values. **e** Confusion matrix with *Lactobacillus* abundance (dominant as positive/deplete as negative) as the reference test and sialidase result (High/Low) as the predicted test. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated.

observed profiles. In previous work, we have observed that the most prevalent vaginal microbiota community compositions in European Caucasian women are those dominated by *L. crispatus* (50%) followed by *L. iners* (25%), whereas the inverse has been reported for Asian and Black women (*L. iners*, 48% and 50% v *L. crispatus*, 26% and 20%, respectively)⁶. Higher frequency of *L. iners* dominated vaginal microbiota has also been reported in Indian⁴³, Karen and Burman⁴⁴, Hispanic and African American pregnant women^{9,10,17}. Our results therefore provide further evidence that ethnic and/or geographical differences are a major source of variation in the underlying structure and composition of the vaginal microbiome during pregnancy.

Correlation analyses in our patient cohort showed that *L. crispatus* was negatively correlated with almost all other vaginal taxa, highlighting its well-described exclusionary behavior in the vaginal niche that appears to be broadly consistent across different ethnic groups^{6,13,17,45}. However, we also observed a strong negative correlation between *L. iners* and most other vaginal taxa. This was surprising given that *L. iners* has been shown to co-occur with various BV-taxa in other ethnic groups including Caucasian and African American women^{9,46}. The highly negative correlation observed between *L. iners* and *L. crispatus* might be attributable to overlapping ecological functions in the vagina^{5,47}. Compared to *L. crispatus*, *L. iners* has a substantially smaller genome, thought to be indicative of a symbiotic and/or parasitic role in the vaginal niche^{5,47,48}. Similar to *G. vaginalis*, *L. iners* can produce pore-forming toxins such as inerolysin and vaginolyisin, which can lead to lysis of host cells and release of carbon sources, particularly during times of nutrient scarcity^{5,47,48}. This behavior may explain the observed positive relationship between *L. iners* and leukocyte presence in our study cohort. Consistent with these findings, a recent study of 83 healthy Chinese pregnant women reported a positive association between leukocyte esterase concentrations and increased *L. iners* levels determined using quantitative real-time PCR⁴⁹. These women were also reported to have increased presence of white blood cells and other non-*Lactobacillus* morphotypes observed under microscopy. In vitro experiments using human vaginal epithelial cells have also demonstrated that *L. iners* stimulates increased pro-inflammatory cytokine production compared to *L. crispatus*^{50,51}. Despite these findings, we did not observe any relationship between *L. iners* dominance or any vaginal microbiota profile and subsequent risk of PTB in this cohort. This is in contrast to previous studies of predominantly Caucasian women by ourselves and others that have reported a relationship between *L. iners* and increased risk of PTB or cervical shortening, which is clinically used as a marker of PTB risk^{6,7,12,52}. Our data are consistent with earlier reports in Hispanic and African American women indicating that *L. iners* dominance is not a risk factor for PTB in these ethnic groups.

In our sub-cohort analysis, high vaginal sialidase activity was strongly associated with *Lactobacillus* depleted microbiota enriched with BV-associated taxa including *Gardnerella*, *Prevotella*, *Atopobium*, and *Megasphaera* species. This finding is consistent with previous studies in both pregnant and nonpregnant women where the production of sialidase by BV-associated taxa, such as *Gardnerella* spp. and *Prevotella* spp., is thought to be important for

pathogenesis^{27,53,54}. In this context, desialylation of glycolipids and/or glycoproteins (e.g. immunoglobulins, cytokines, cellular receptors, mucins, and antimicrobial molecules) can decrease the ability of host defence responses to recognize and bind to microbes whilst increasing bacterial adherence, invasion, and tissue breakdown^{27,30,54,55}. Interestingly, high sialidase activity was observed in some *Lactobacillus*-dominated vaginal microbiota. These often contained low relative abundance of BV-associated taxa indicating that perhaps even low levels of sialidase-producing bacteria may be sufficient to produce high sialidase activity. It is also possible that sialidase was produced by other microorganisms including viruses and fungi that were not assessed in our study^{56,57}. In contrast to these findings, a proportion of samples (23%) harboring *Lactobacillus* spp. depleted vaginal microbiota enriched with BV-associated taxa did not have high sialidase activity. This was supported by our findings indicating sialidase activity has high specificity (99%) but low sensitivity (14%) in predicting high-diversity, *Lactobacillus* spp. depleted vaginal microbiota. Low sialidase activity in these samples could be due to several factors: 1) the production of sialidase by BV-taxa, including *Gardnerella* spp., is strain-dependent^{27,30,31} 2) binding of sialidase to the BV test kit substrate may be influenced by vaginal pH^{26,58} 3) the amount of sialidase produced by BV-taxa maybe insufficient to reach the detection limit of the BV test kit; and/or 4) the bacterial load of BV-taxa, which was not quantified in our study, may be inadequate to produce sufficient sialidase for detection by the BV test kit. Overall, our findings suggest that high sialidase activity is significantly associated with BV-taxa but cannot be used as an accurate proxy for high-diversity, *Lactobacillus* spp. depleted vaginal microbiota in pregnant women.

Elevated sialidase activity has previously been associated with spontaneous PTB and late miscarriage in a North American study of 1806 pregnant women when measured at 12 weeks gestation³⁰, and with early PTB and low birth weight in a mid-trimester study of 579 Danish women⁵⁹. Additionally, in a recent study of 85 Chinese women, sialidase activity was associated with subsequent cervical cerclage failure, a risk factor for PTB³¹. However, in our study, no significant association was found between elevated sialidase activity and PTB. Our findings are similar to a previous nested case-control study of 126 pregnant women, where elevated sialidase activity measured at mid-pregnancy was not associated with PTB risk⁵³. The observed differences between studies could be due to different methods for measuring sialidase activity where a single cut-off was used rather than a quantitative measurement. An association between vaginal leukocytes and preterm labor or risk factors of PTB including reproductive tract infections, histologic chorioamnionitis and inflammatory cytokines have also previously been reported^{33,34,39,60–62}. Here, we classified leukocyte presence based on vaginal wet mount microscopy, which has been shown to be useful for identification of cervicovaginal inflammation in pregnant women^{39,63}. Logistic regression analyses indicated that neither leukocyte presence, sialidase activity nor vaginal microbiota contributed to preterm risk in our study cohort. The rate of PTB in our study population was 5.4% which is slightly lower than the background rate in China (7.2%)⁶⁴, but comparable to published preterm birth rates

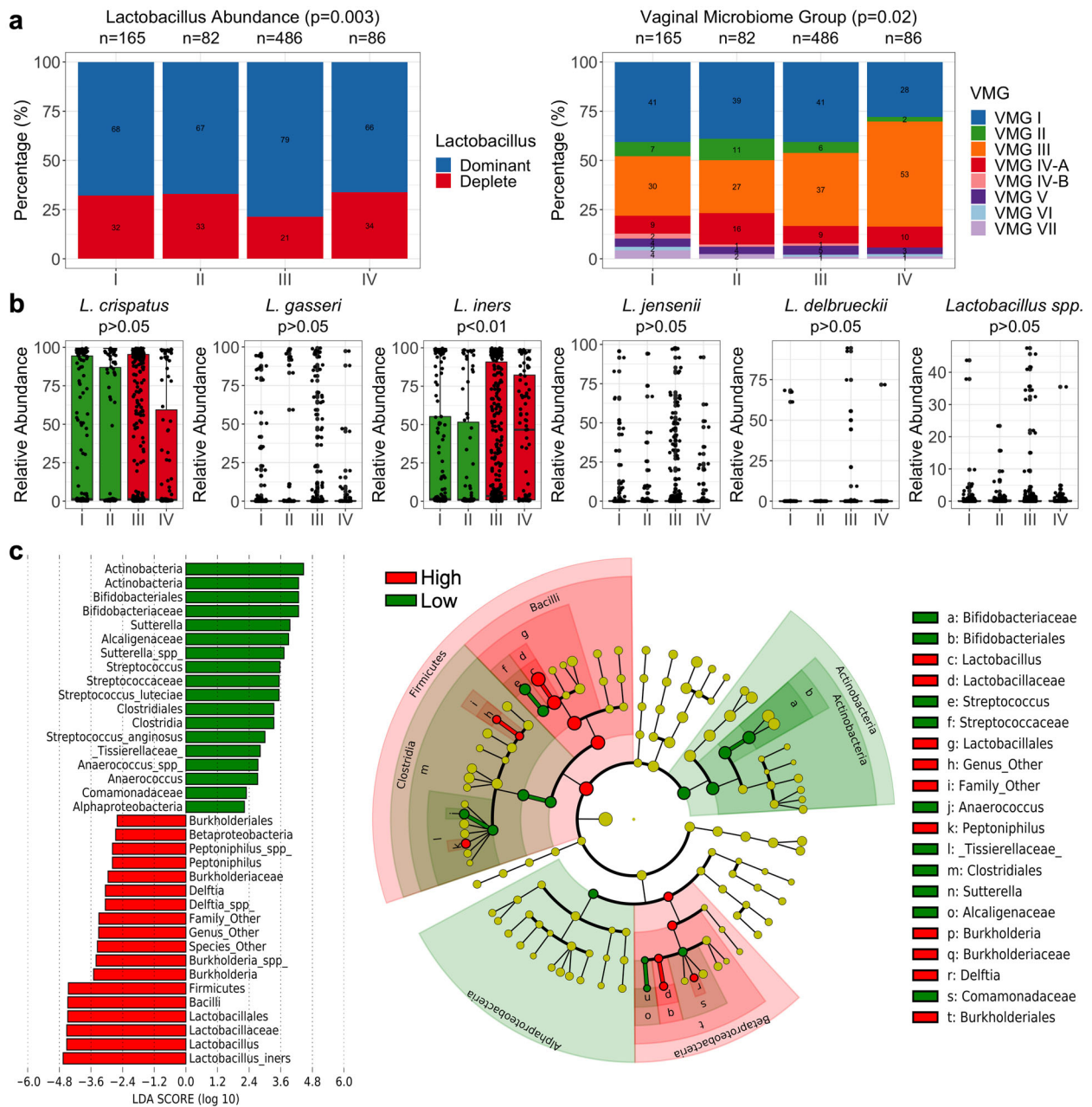


Fig. 3 Vaginal microbiome composition and leukocyte presence. **a** Stacked bar showing *Lactobacillus* abundance groups and vaginal microbiome groups (VMG) of women with leukocyte grades of I (0–5 counts/Hp), II (5–15 counts/Hp), III (15–30 counts/Hp) and IV (>30 counts/Hp). **b** Boxplots showing relative abundance of *L. crispatus*, *L. gasseri*, *L. iners*, *L. jensenii*, *L. delbrueckii*, and *Lactobacillus* spp. in women with leukocyte low (grades I and II, green) and high (grades III and IV, red). The bounds of the box represent the first and third quartiles, center line represents the median, and whiskers show min-to-max values. **c** LDA of differentially abundant taxa and cladogram displaying bacterial clades and nodes identified as differentially abundant from LDA analysis in vaginal microbiota of women with leukocyte low (grades I–II, green) and high (grades III–IV, red). Statistical significance of categorical variables based on Fisher's Exact Test and continuous variables based on Kruskal-Wallis Test.

for the local population in Guangzhou (4.15% and 6.0%)^{65–67}. These rates are considerably less than the global average of ~10%⁶⁸. Our findings therefore highlight differences in the vaginal microbiota and/or host immune response as a potential contributing factor to reduced incidence of PTB in Chinese women.

There are several limitations to our study that should be considered when interpreting the results. Birth outcome data was unavailable for 48% of the women in our cohort. These women did not return to the recruitment site of Nanfang Hospital and due to data protection considerations, were unable to be contacted for follow-up. Although vaginal microbiota composition in these

women was comparable to those with available outcome data, non-observed outcomes may represent a potential source of bias. Further, the pregnancy vaginal microbiota can be impacted by a variety of factors including pregnancy history, smoking status, genetics and hygiene habits^{21,69,70}. This information was not available for our patient cohort and cross-sectional sampling of the vaginal microbiota did not permit us to assess the potential role of vaginal microbiota dynamics in shaping clinical outcomes^{17,71,72}. Our study was also limited to the assessment of the bacterial component of the vaginal microbiota via 16S rRNA gene sequencing, which does not provide sufficient resolution to

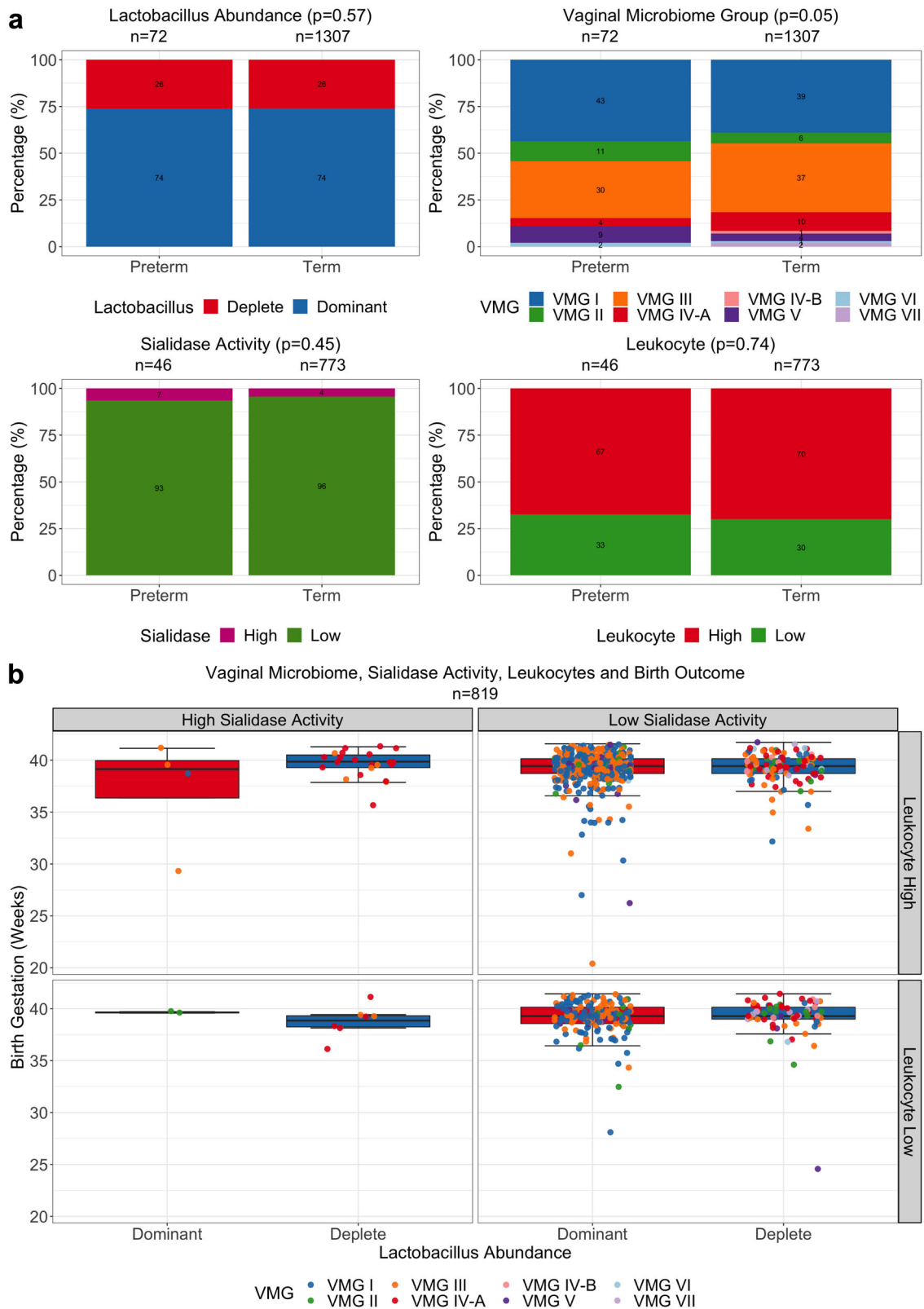


Fig. 4 Association between the vaginal microbiome, sialidase activity and leukocyte presence with birth outcome. **a** Stacked bar plots showing proportion of *Lactobacillus* abundance (deplete, red; or dominant, blue), vaginal microbiome groups (VMG), sialidase activity (high, pink; or low, red) and leukocyte low (grades I-II, green) or high (grades III-IV, red) in preterm and term groups. Statistical significance based on Fisher's Exact Test. **b** Boxplots showing birth gestation in weeks grouped by *Lactobacillus* abundance, sialidase activity, and leukocyte presence or absence with dot points colored by VMG. The bounds of the box represent the first and third quartiles, center line represents the median, and whiskers show min-to-max values. No significant difference between boxplots with $n \geq 3$ based on Kruskal-Wallis Test, $p > 0.05$.

account for strain-level variations associated with potential function of microbial communities^{73,74}. Future studies could apply function-based profiling approaches to enable a deeper insight into potential relationships between predicted function of microbiome communities within different ethnic groups and preterm birth risk phenotype.

In conclusion, our study provides new insight into the vaginal microbial composition and structure of Chinese pregnant women. Although sialidase activity was predictive of high diversity vaginal microbiota compositions, neither were associated with increased risk of PTB. *Lactobacillus* spp. depleted vaginal microbiomes with high sialidase activity and leukocyte presence are not associated with higher risk of PTB in Chinese women. Our results provide further evidence that ethnicity is an important determinant of microbiota-host interactions during pregnancy and highlight the need for further investigations into the mechanisms underpinning these relationships.

METHODS

Study design

This prospective study was reviewed and approved ([2013] EC (100)) by the Ethical Committee of Nanfang Hospital, Southern Medical University, Guangzhou, China. Written, informed consent was obtained from all participants. Pregnant women at their first prenatal visit were recruited from the outpatient clinic of Nanfang Hospital of Southern Medical University in Guangzhou, China, from January 2015 to December 2018. All women were recruited consecutively, with no disruption to recruitment during this period. Women who received antibiotics, prebiotics or probiotics within 30 days prior to vaginal swab collection and/or who had sexual activity within 48 h of sample collection were excluded. Metadata collected included maternal age, mode of labor, gestation at delivery in days (calculated using the last menstrual period and/or ultrasound data) and chorioamnionitis (Supplementary Data 1).

Sample collection and processing

Vaginal samples were collected using a sterile swab that was inserted into the vaginal posterior fornix, before being gently rotated 360° for approximately 20 rotations prior to removal. All swabs were placed immediately on ice and were then stored at -20 °C within 1 h of collection before being transferred to -80 °C within 24 h for long-term storage. For DNA extraction, swabs were immersed and vigorously mixed in 500 µL of sterile water before the solution was transferred to a clean 2 mL centrifuge tube. The sample was then vortexed for 5 sec before being centrifuged at 13,800 g for 10 min. The supernatant was removed, and the pellet retained for DNA extraction. For PCR negative controls ($n=82$), 2 µL of diethylpyrocarbonate (DEPC) water was used as non-template control. For DNA extraction kit negative controls ($n=21$), 500 µL of sterile water was used instead of test samples for the DNA extraction process.

DNA extraction and 16S rRNA (V4) amplicon sequencing

DNA was extracted from vaginal swabs manually using the BioTeke bacterial DNA extraction kit (BioTeke Corporation, Cat #DP7001) per manufacturer's instructions ($n=1176$) or via an automated protocol using the Bioeasy bacterial DNA extraction kit (Shenzhen Bioeasy Biotechnology Co. Ltd., Cat #YRMBN7001) on a Thermo Scientific™ KingFisher™ Flex Purification System ($n=1620$). The V4 region of bacterial 16S rRNA was amplified using barcoded V4F 5' GTGCCAGCMGCCGCGTAA 3' forward and V4R 5' CTACCNNGGTATCTAAT 3' reverse primers. The PCR condition included an initial denaturation step at 94 °C for 5 min, 30 cycles of 94 °C for 30 sec, 52 °C for 30 sec, 72 °C for 45 sec and a final extension step at 72 °C for 5 min. Each 20 µL reaction volume consisted of 10 µL of AceQqPCR SYBR Green Master Mix (Vazyme, Nanjing, China), 0.4 µL prime V4F, 0.4 µL primer V4R, 0.4 µL ROX Reference Dye 2, 2 µL template DNA, 8.8 µL sterilized distilled water. Equimolar amplicon suspensions were combined and subjected to paired-end 101 bp sequencing on an Illumina HiSeq sequencer at the Beijing Genomics Institute (BGI; Beijing, China).

Sialidase enzyme activity detection

A total of 848 women had an additional vaginal swab collected for detection of sialidase enzyme activity. Sialidase enzyme activity was measured as per manufacturer's instructions using a single-enzyme BV kit based on the colorimetric method (Zhuhai DL Biotech. Co. Ltd). Briefly, each swab sample was immersed into a BV test bottle solution containing 5-bromo-4-chloro-3-indolyl- α -D-N-acetylneuraminic acid (BCIN) as a substrate⁷⁵. The BCIN substrate hydrolyzes and reacts with the added 1 – 2 drops of BV chromogenic solution containing hydroxide and potassium acetate, resulting in a color reaction that was measured using the BV-10 analyser (Zhuhai DL Biotech. Co. Ltd). The resulting sialidase activity is defined based on enzyme unit (U), where one U is the amount of sialidase required to catalyze 1 nmol of BCIN per minute in a ml. The amount of sialidase activity corresponded to color changes in the BV test bottle with ≥ 7.8 U/ml for a positive result (blue or green) and 0 – 7.8U/ml for a negative result (yellow). The samples were defined as having high or low sialidase activity based on ≥ 7.8 U/ml or 0 – 7.8U/ml, respectively⁷⁶.

Leukocyte wet mount

Preparation of leukocyte wet mounts was performed by agitating vaginal swabs in 1 mL of saline solution and a single drop of the fluid was then placed on a glass slide^{76,77}. A cover slip was applied on the droplet and the slides were then inspected under 400x magnification. The average white blood cell count (non-clumped) in 10 nonadjacent fields of magnificant were determined. Leukocyte counts were categorized into grades I (0–5 counts/Hp), II (5–10 counts/Hp), III (10–15 counts/Hp) or IV (>30 counts/Hp). Grades I and II were further classified as leukocyte low (<15 counts/Hp) and grades III–IV as high (>15 counts/Hp)⁶³.

Chorioamnionitis diagnosis

Chorioamnionitis was diagnosed using a combination of clinical symptoms and histological assessment. Clinical symptoms included the presence of high maternal fever (≥ 39 °C), maternal tachycardia (>100 beats/min) and malodorous amniotic fluid. Histological chorioamnionitis was diagnosed based on the presence of inflammatory cells in the chorionic plate and/or the chorioamniotic membranes consistent with acute inflammation and the presence of neutrophils in the wall of the umbilical vessels and Wharton's jelly^{78,79}. Staging and grading of chorioamnionitis was performed following the guidelines of the Amsterdam Placental Workshop Group Consensus Statement⁸⁰.

Bioinformatics and statistical analyses

The raw sequencing reads were processed using the DADA2 package (v1.6.0) in R (v3.4.3)⁸¹. The raw paired-end reads were assigned to samples based on their unique barcodes and truncated by cutting off the barcodes and primer sequences. The reads were then processed using the DADA2 package (v1.6.0) in R (v3.4.3) according to the following steps: quality controls, dereplication, error rate calculations, sequences denoising, paired-end reads merging, amplicon sequence variant (ASV) table constructing and removal of chimeras. The sequences in the ASV table were annotated using the RDP Classifier⁸² with the GreenGenes database⁸³. The sequences assigned to *Lactobacillus* were further classified into species level using UCLUST⁸⁴ with the customized *Lactobacillus* database (V4 region extracted from representative sequences of *Lactobacillus* species from RDP database).

All downstream bioinformatic and statistical analyses were performed in R (v4.0.0)⁸⁵ unless otherwise stated. Low abundance taxa accounting for less than 0.1% within-sample relative abundance and present in less than 1% of samples were filtered using phyloseq (v1.32.0)⁸⁶ and genefilter (v1.70.0)⁸⁷ packages. This left a total of 167 ASVs representing 94 unique taxa. Extraction kit and reagent negative controls ($n=103$) were used to identify potential contaminants using the decontam package (v1.9.0) with a prevalence threshold of 0.1 and 0.5 respectively⁸⁸. A total of 12 contaminants were assigned as *Pseudomonas veronii*, *Bifidobacterium adolescentis*, *Klebsiella* spp., *Enhydrobacter* spp., *Chryseobacterium* spp., *Pedobacter* spp., *Bradyrhizobium* spp., *Methylobacterium* spp., *Escherichia coli*, *Acinetobacter johnsonii* and *Sneathia* spp., and *Species_Other*. Samples with a library size of <2000 reads were excluded from the analysis ($n=90$) (Supplementary Fig. 1).

Each vaginal bacterial profile in the cohort was initially classified on the basis of the highest relative abundance of the dominant species taxon within the sample^{2,89}. Dominant species with <30% abundance were

classified as "Other" ($n=16$) and a single sample dominated by *Mesorhizobium* spp. was classified as "Outlier". These samples were excluded from subsequent analyses. A total of 19 dominant (>30% relative abundance) species groups were then classified into VMG consisting of VMG I (*L. crispatus*); II (*L. gasseri*); III (*L. iners*); IV-A (BV-associated species; *Atopobium vaginae*, *Gardnerella* spp., *Megasphaera* spp., *Prevotella* spp., *Shuttleworthia* spp., *Veillonella* spp.)^{40,69}; IV-B (Non-BV taxa; *Aerococcus* spp., *Anaerococcus* spp., *Streptococcus agalactiae*, *Streptococcus anginosus*, *Streptococcus luteiae*, *Streptococcus* spp.)⁹⁰; V (*L. jensenii*); VI (Other *Lactobacillus*; *L. delbrueckii*, *Lactobacillus* spp.) and VII (*Bifidobacterium* spp.). Analyses were also performed on samples classified into *Lactobacillus* species dominated and depleted groups based on *Lactobacillus* abundance of greater or lower than 90%⁹¹.

Data were visualized using ggplot2 package (v3.3.1)⁹². Statistical difference for continuous variables between groups were determined using either Mann-Whitney Test with Bonferroni p-adjusted values or Kruskal-Wallis Test with pairwise comparison between groups and Bonferroni p-adjusted values. Association between categorical groups were determined using Fisher's Exact Test. For all statistical tests, significance was based on $p < 0.05$ unless otherwise stated. Spearman's correlation plots were generated from relative abundance using corplot package (v0.84) with ward.D2 agglomerative method⁹³. The Shannon diversity index and richness (number of species observed) were computed using the vegan package (v2.5–6)⁹⁴, the former calculated for both raw and normalized read counts using the trimmed mean of M-values normalization method in the edgeR package (v3.30.3)⁹⁵. LefSe analysis was performed using per-sample normalization to sum of values to 1 M with a Linear Discriminant Analysis (LDA) threshold of 2.0 for discriminative features and an alpha of 0.05 for factorial Kruskal-Wallis test among classes and alpha 0.05 for pairwise Wilcoxon test between subclasses⁹⁶. The caret package (v6.0–86)⁹⁷ was used to compute confusion matrices with a prevalence of 0.15 and perform multiple logistic regression using gestation at delivery (days) as a binary outcome (Preterm, <259 days and Term, ≥259 days) for the response variable and *Lactobacillus* abundance, sialidase activity and leukocyte wet mount results as explanatory variables.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

Raw data files for the sequence data used in this study are publicly available through the European Nucleotide Archive (<https://www.ebi.ac.uk/ena>) under accession numbers (PRJNA706523).

CODE AVAILABILITY

Key analysis code and processed datasets are publicly available the GitHub repository; https://github.com/sherianne/NgChen_2021_ChinaPregnancyCohort.

Received: 14 June 2021; Accepted: 29 November 2021;

Published online: 20 December 2021

REFERENCES

- 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).
- Fettweis, J. M. et al. The vaginal microbiome and preterm birth. *Nat. Med* **25**, 1012–1021 (2019).
- Kindinger, L. M. et al. Relationship between vaginal microbial dysbiosis, inflammation, and pregnancy outcomes in cervical cerclage. *Sci. Transl. Med* **8**, 350ra102 (2016).
- Amabebe, E. & Anumba, D. O. C. The vaginal microenvironment: the physiologic role of *Lactobacilli*. *Front Med (Lausanne)* **5**, 181 (2018).
- France, M. T., Mendes-Soares, H., Forney, L. J. & Schloss, P. D. Genomic Comparisons of *Lactobacillus crispatus* and *Lactobacillus iners* Reveal Potential Ecological Drivers of Community Composition in the Vagina. *Appl. Environ. Microbiol.* **82**, 7063–7073 (2016).
- Kindinger, L. M. et al. The interaction between vaginal microbiota, cervical length, and vaginal progesterone treatment for preterm birth risk. *Microbiome* **5**, 6 (2017).

- Payne, M. S. et al. A specific bacterial DNA signature in the vagina of Australian women in midpregnancy predicts high risk of spontaneous preterm birth (the Predict1000 study). *Am J Obstet Gynecol*, <https://doi.org/10.1016/j.ajog.2020.08.034> (2020).
- Stafford, G. P. et al. Spontaneous Preterm Birth Is Associated with Differential Expression of Vaginal Metabolites by *Lactobacilli*-Dominated Microflora. *Front Physiol.* **8**, 615 (2017).
- 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. USA* **114**, 9966–9971 (2017).
- Elovitz, M. A. et al. Cervicovaginal microbiota and local immune response modulate the risk of spontaneous preterm delivery. *Nat. Commun.* **10**, 1305 (2019).
- Kindinger, L. M. et al. The effect of gestational age and cervical length measurements in the prediction of spontaneous preterm birth in twin pregnancies: an individual patient level meta-analysis. *BJOG* **123**, 877–884 (2016).
- Petricevic, L. et al. Characterisation of the vaginal *Lactobacillus* microbiota associated with preterm delivery. *Sci. Rep.* **4**, 5136 (2014).
- Brown, R. G. et al. Establishment of vaginal microbiota composition in early pregnancy and its association with subsequent preterm prelabor rupture of the fetal membranes. *Transl. Res.* **207**, 30–43 (2019).
- DiGiulio, D. B. et al. Temporal and spatial variation of the human microbiota during pregnancy. *Proc. Natl Acad. Sci. USA* **112**, 11060–11065 (2015).
- Freitas, A. C., Bocking, A., Hill, J. E., Money, D. M. & Group, V. R. Increased richness and diversity of the vaginal microbiota and spontaneous preterm birth. *Microbiome* **6**, 117 (2018).
- Blencowe, H. et al. Born too soon: the global epidemiology of 15 million preterm births. *Reprod. Health* **10**(Suppl 1), S2 (2013).
- Serrano, M. G. et al. Racioethnic diversity in the dynamics of the vaginal microbiome during pregnancy. *Nat. Med* **25**, 1001–1011 (2019).
- Amsel, R. et al. Nonspecific vaginitis: diagnostic criteria and microbial and epidemiologic associations. *Am. J. Med* **74**, 14–22 (1983).
- Nugent, R. P., Krohn, M. A. & Hillier, S. L. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of Gram stain interpretation. *J. Clin. Microbiol.* **29**, 297–301 (1991).
- Donders, G. G. et al. Predictive value for preterm birth of abnormal vaginal flora, bacterial vaginosis and aerobic vaginitis during the first trimester of pregnancy. *BJOG: Int. J. Obstet. Gynaecol.* **116**, 1315–1324 (2009).
- Guaschino, S., De Seta, F., Piccoli, M., Maso, G. & Alberico, S. Aetiology of preterm labour: bacterial vaginosis. *BJOG* **113**(Suppl 3), 46–51 (2006).
- Leitch, H. et al. Bacterial vaginosis as a risk factor for preterm delivery: a meta-analysis. *Am. J. Obstet. Gynecol.* **189**, 139–147 (2003).
- Hillier, S. L. et al. Association between bacterial vaginosis and preterm delivery of a low-birth-weight infant. The Vaginal Infections and Prematurity Study Group. *N. Engl. J. Med* **333**, 1737–1742 (1995).
- Colonna, C. & Steelman, M. in *StatPearls* (2021).
- Madhivanan, P. et al. Performance of BVBlue rapid test in detecting bacterial vaginosis among women in Mysore, India. *Infect. Dis. Obstet. Gynecol.* **2014**, 908313 (2014).
- Myziuk, L., Romanowski, B. & Johnson, S. C. BVBlue test for diagnosis of bacterial vaginosis. *J. Clin. Microbiol.* **41**, 1925–1928 (2003).
- Briselden, A. M., Moncla, B. J., Stevens, C. E. & Hillier, S. L. Sialidases (Neuraminidases) in bacterial vaginosis and bacterial vaginosis-associated microflora. *J. Clin. Microbiol.* **30**, 663–666 (1992).
- Hardy, L. et al. The presence of the putative *Gardnerella vaginalis* sialidase A gene in vaginal specimens is associated with bacterial vaginosis biofilm. *PLoS One* **12**, e0172522 (2017).
- Jung, H. S., Ehlers, M. M., Lombaard, H., Redelinghuys, M. J. & Kock, M. M. Etiology of bacterial vaginosis and polymicrobial biofilm formation. *Crit. Rev. Microbiol.* **43**, 651–667 (2017).
- Cauci, S. & Culhane, J. F. High sialidase levels increase preterm birth risk among women who are bacterial vaginosis-positive in early gestation. *Am. J. Obstet. Gynecol.* **204**, 142 e141–142 e149 (2011).
- Fang, J., Chen, L., Chen, Z., Jiang, X. & Pan, M. Association of the vaginal microbiota with pregnancy outcomes in Chinese women after cervical cerclage. *Reprod. Biomed. Online* **41**, 698–706 (2020).
- Goldenberg, R. L., Hauth, J. C. & Andrews, W. W. Intrauterine infection and preterm delivery. *N. Engl. J. Med* **342**, 1500–1507 (2000).
- Yudin, M. H. et al. Vaginal polymorphonuclear leukocytes and bacterial vaginosis as markers for histologic endometritis among women without symptoms of pelvic inflammatory disease. *Am. J. Obstet. Gynecol.* **188**, 318–323 (2003).
- Geisler, W. M., Yu, S., Venglarik, M. & Schwabke, J. R. Vaginal leukocyte counts in women with bacterial vaginosis: relation to vaginal and cervical infections. *Sex. Transm. Infect.* **80**, 401–405 (2004).

35. Hakakha, M. Leukorrhea and bacterial vaginosis as in-office predictors of cervical infection in high-risk women. *Obstet. Gynecol.* **100**, 808–812 (2002).
36. Ramsey, P. S. et al. Use of vaginal polymorphonuclear to epithelial cell ratios for the prediction of preterm birth. *Obstet. Gynecol.* **105**, 139–144 (2005).
37. Verstraelen, H. et al. Modified classification of Gram-stained vaginal smears to predict spontaneous preterm birth: a prospective cohort study. *Am. J. Obstet. Gynecol.* **196**, 528 e521–526 (2007).
38. Simhan, H. N., Caritis, S. N., Krohn, M. A. & Hillier, S. L. Elevated vaginal pH and neutrophils are associated strongly with early spontaneous preterm birth. *Am. J. Obstet. Gynecol.* **189**, 1150–1154 (2003).
39. Nenadić, D. B. & Pavlović, M. D. Cervical fluid cytokines in pregnant women: Relation to vaginal wet mount findings and polymorphonuclear leukocyte counts. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **140**, 165–170 (2008).
40. Gajer, P. et al. Temporal Dynamics of the Human Vaginal Microbiota. *Sci. Transl. Med.* **4**, 132ra52 (2013).
41. Stout, M. J. et al. Early pregnancy vaginal microbiome trends and preterm birth. *Am. J. Obstet. Gynecol.* **217**, 356 e351–356 e318 (2017).
42. He, Y., Huang, Y., Zhang, Z., Yu, F. & Zheng, Y. Exploring profile and potential influencers of vaginal microbiome among asymptomatic pregnant Chinese women. *PeerJ* **7**, e8172 (2019).
43. Mehta, O. et al. Vaginal microbiome of pregnant Indian women: insights into the genome of dominant *Lactobacillus* species. *Micro. Ecol.* **80**, 487–499 (2020).
44. Kumar, M. et al. Vaginal microbiota and cytokine levels predict preterm delivery in Asian women. *Front Cell Infect. Microbiol* **11**, 639665 (2021).
45. Ravel, J. et al. Vaginal microbiome of reproductive-age women. *Proc. Natl Acad. Sci. USA* **108**(Suppl 1), 4680–4687 (2011).
46. Verstraelen, H. et al. Longitudinal analysis of the vaginal microflora in pregnancy suggests that *L. crispatus* promotes the stability of the normal vaginal microflora and that *L. gasseri* and/or *L. iners* are more conducive to the occurrence of abnormal vaginal microflora. *BMC Microbiology* **9**, 116 (2009).
47. Witkin, S. S., Moron, A. F., Linhares, I. M. & Forney, L. J. Influence of *Lactobacillus crispatus*, *Lactobacillus iners* and *Gardnerella vaginalis* on bacterial vaginal composition in pregnant women. *Arch Gynecol Obstet*, <https://doi.org/10.1007/s00404-021-05978-z> (2021).
48. Petrova, M. I., Reid, G., Vaneechoutte, M. & Lebeer, S. *Lactobacillus iners*: friend or foe? *Trends Microbiol* **25**, 182–191 (2017).
49. Zheng, N. et al. *Lactobacillus iners* is associated with vaginal dysbiosis in healthy pregnant women: a preliminary Study. *Biomed. Res Int* **2019**, 6079734 (2019).
50. Anahtar, MelisN. et al. Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. *Immunity* **42**, 965–976 (2015).
51. 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**, 1989–1999 (2014).
52. Witkin, S. S. et al. Vaginal Biomarkers That Predict Cervical Length and Dominant Bacteria in the Vaginal Microbiomes of Pregnant Women. *mBio* **10**, 02242–19 (2019).
53. Andrews, W. W. et al. The Preterm Prediction Study: Failure of midtrimester cervical sialidase level elevation to predict subsequent spontaneous preterm birth. *Am. J. Obstet. Gynecol.* **180**, 1151–1154 (1999).
54. Cauci, S. et al. Immunoglobulin A response against *Gardnerella vaginalis* hemolysin and sialidase activity in bacterial vaginosis. *Am. J. Obstet. Gynecol.* **178**, 511–515 (1998).
55. Cauci, S. et al. Determination of immunoglobulin A against *Gardnerella vaginalis* hemolysin, sialidase, and proliadase activities in vaginal fluid: implications for adverse pregnancy outcomes. *J. Clin. Microbiol.* **41**, 435–438 (2003).
56. Schwerdtfeger, S. M. & Melzig, M. F. Sialidases in biological systems. *Pharmazie* **65**, 551–561 (2010).
57. Taylor, G. Sialidases: structures, biological significance and therapeutic potential. *Curr. Opin. Struct. Biol.* **6**, 830–837 (1996).
58. Agarwal, K. & Lewis, A. L. Vaginal sialoglycan foraging by *Gardnerella vaginalis*: Mucus barriers as a meal for unwelcome guests? *Glycobiology*, <https://doi.org/10.1093/glycob/cwab024> (2021).
59. Cauci, S., McGregor, J., Thorsen, P., Grove, J. & Guaschino, S. Combination of vaginal pH with vaginal sialidase and proliadase activities for prediction of low birth weight and preterm birth. *Am. J. Obstet. Gynecol.* **192**, 489–496 (2005).
60. Matsubara, S., Yamada, T., Minakami, H. & Sato, I. Stimulated polymorphonuclear leukocytes in vaginal secretions from patients with preterm labor. *Gynecol. Obstet. Invest* **45**, 35–40 (1998).
61. Yamada, T. et al. Relation between viability of vaginal polymorphonuclear leukocytes and presence of histologic chorioamnionitis. *Acta Obstet. Gynecol. Scand.* **79**, 818–823 (2000).
62. Yamada, T., Minakami, H., Matsubara, S., Yatsuda, T. & Sato, I. Changes in polymorphonuclear leukocytes in the vagina of patients with preterm labor. *Gynecol. Obstet. Invest* **45**, 32–34 (1998).
63. Zhang, J. et al. Computerized detection of leukocytes in microscopic leukorrhea images. *Med Phys.* **44**, 4620–4629 (2017).
64. Guo, T. et al. The association between ambient temperature and the risk of preterm birth in China. *Sci. Total Environ.* **613–614**, 439–446 (2018).
65. He, J. R. et al. Ambient Temperature and the Risk of Preterm Birth in Guangzhou, China (2001–2011). *Environ. Health Perspect.* **124**, 1100–1106 (2016).
66. Lu, J. et al. Increasing trends in incidence of preterm birth among 2.5 million newborns in Guangzhou, China, 2001 to 2016: an age-period-cohort analysis. *BMC Public Health* **20**, 1653 (2020).
67. Miao, H. et al. Adverse birth outcomes in Guangdong province, China, 2014–2017: a spatiotemporal analysis of 2.9 million births. *BMJ Open* **9**, e030629 (2019).
68. Chawanpaiboon, S. et al. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. *Lancet Glob. Health* **7**, e37–e46 (2019).
69. Onderdonk, A. B., Delaney, M. L. & Fichorova, R. N. The human microbiome during bacterial vaginosis. *Clin. Microbiol Rev.* **29**, 223–238 (2016).
70. Turovskiy, Y., Sutyak Noll, K. & Chikindas, M. L. The aetiology of bacterial vaginosis. *J. Appl Microbiol* **110**, 1105–1128 (2011).
71. MacIntyre, D. A. et al. The vaginal microbiome during pregnancy and the post-partum period in a European population. *Sci. Rep.* **5**, 8988 (2015).
72. 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).
73. Douglas, G. M. et al. Overcoming barriers to early disease intervention. *Nat. Biotechnol.* **38**, 669–688 (2020).
74. Langille, M. G. et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **31**, 814–821 (2013).
75. Wiggins, R. et al. Use of 5-bromo-4-chloro-3-indolyl- α -D-N-acetylneuraminic acid in a novel spot test to identify sialidase activity in vaginal swabs from women with bacterial vaginosis. *J. Clin. Microbiol* **38**, 3096–3097 (2000).
76. Kottmel, A. & Petersen, E. E. Vaginal wet mount. *J. Sex. Med* **10**, 2616–2619 (2013).
77. Galinanes, S. et al. Vaginal inflammatory status in pregnant women with normal and pathogenic microbiota in lower genital tract. *ISRN Obstet. Gynecol.* **2011**, 835926 (2011).
78. Tita, A. T. & Andrews, W. W. Diagnosis and management of clinical chorioamnionitis. *Clin. Perinatol.* **37**, 339–354 (2010).
79. Yoon, B. H. et al. Clinical significance of intra-amniotic inflammation in patients with preterm labor and intact membranes. *Am. J. Obstet. Gynecol.* **185**, 1130–1136 (2001).
80. Khong, T. Y. et al. Sampling and Definitions of Placental Lesions: Amsterdam Placental Workshop Group Consensus Statement. *Arch. Pathol. Lab Med* **140**, 698–713 (2016).
81. Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
82. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ. Microbiol* **73**, 5261–5267 (2007).
83. DeSantis, T. Z. et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ. Microbiol* **72**, 5069–5072 (2006).
84. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
85. R Development Core Team, R. R: *A Language and Environment for Statistical Computing*. (2011).
86. McMurdie, P. J. & Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **8**, 0061217 (2013).
87. genefilter: methods for filtering genes from high-throughput experiments (2020).
88. Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 1–14 (2018).
89. Brooks, J. P. et al. Changes in vaginal community state types reflect major shifts in the microbiome. *Micro. Ecol. Health Dis.* **28**, 1303265 (2017).
90. van de Wijgert, J. H. H. M. et al. Pathobionts in the Vaginal Microbiota: Individual Participant Data Meta-Analysis of Three Sequencing Studies. *Front. Cell. Infect. Microbiol.* **10**, 00129 (2020).
91. Fu, M. et al. Alterations in vaginal microbiota and associated metabolome in women with recurrent implantation failure. *mBio* **11**, 03242–19 (2020).
92. Wickham, H. *ggplot2: Elegant graphics for data analysis*. (Springer Verlag, 2009).
93. corplot: Visualization of a correlation matrix v. 0.84 (2017).
94. vegan: Community ecology package v. 2.5-6 (2019).
95. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
96. Segata, N. et al. Metagenomic biomarker discovery and explanation. *Genome Biology* **12**, r60 (2011).
97. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Software* **28**, i05 (2008).

ACKNOWLEDGEMENTS

We would like to thank all patients who have participated in this study. This work was funded by National Natural Science Foundation of China (NSFC81925062, 82002201), National Key R&D Program of China (2019YFA0802300) and the March of Dimes European Prematurity Research Centre at Imperial College London.

AUTHOR CONTRIBUTIONS

S.N. and M.C. contributed equally to this study. S.N., M.C., D.A.M., and H.Z. conceived and designed the study. Clinical sampling and coordination of metadata were performed by M.C., X.W., Z.D.Z., H.S., Y.H., and Y.W. Experiments were performed by M.C., Z.Y., and W.Q. Data processing and analysis were performed S.N., M.C., and D.A.M. S.N. and M.C. prepared all figures and tables. S.N. and D.A.M. wrote the first draft of the manuscript. All authors critically reviewed, read, and approved the final manuscript.

COMPETING INTERESTS

The authors declare that there are no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41522-021-00261-0>.

Correspondence and requests for materials should be addressed to David A. MacIntyre or Hongwei Zhou.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021