## Supplemental Figures

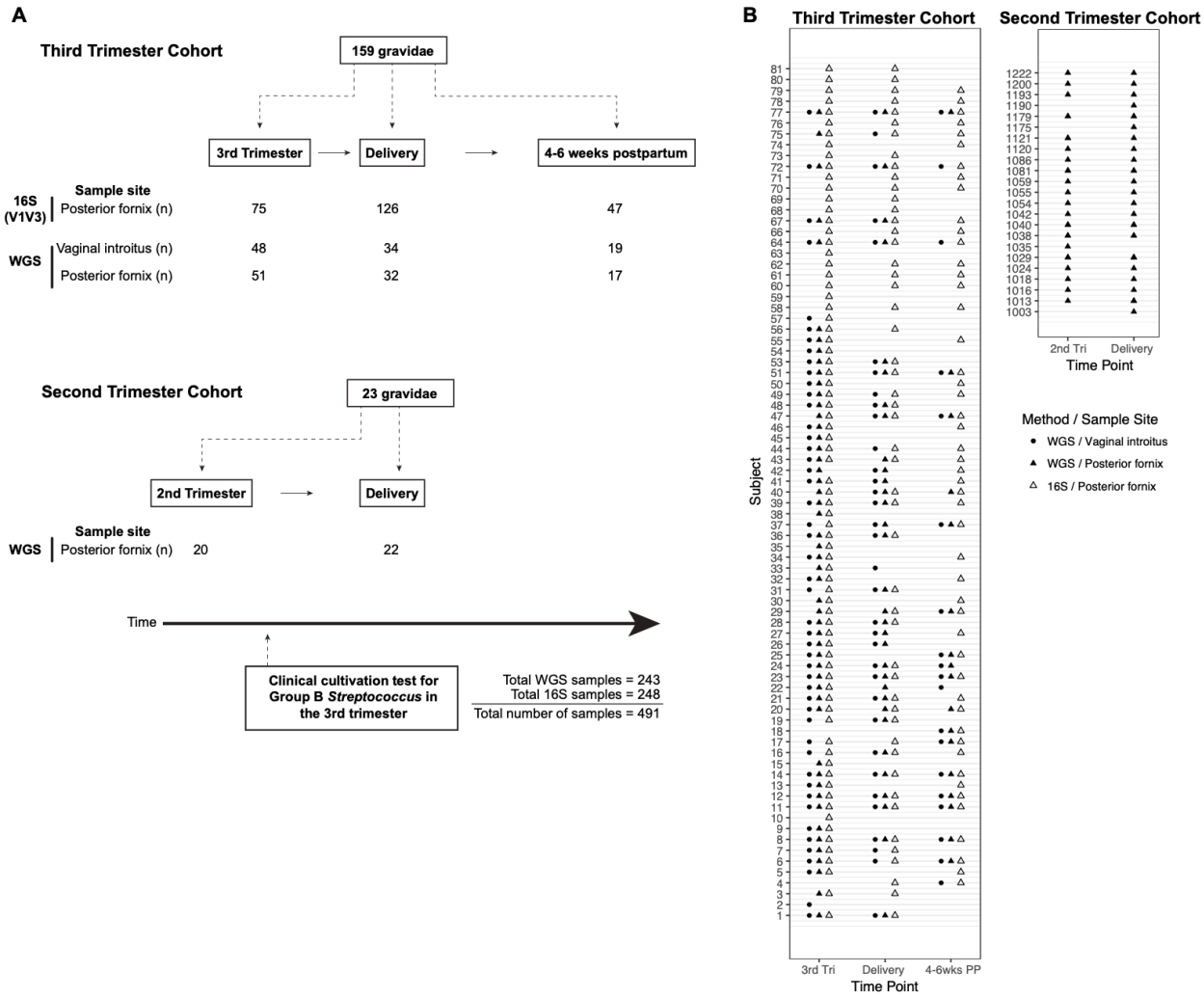



**Figure S1.** Sampling timeline and regimen from among a total of n=182 gravidae, Related to Figure 1**.** (**A**) Sample timeline and sample numbers. 102 gravidae were enrolled longitudinally and sampled in the second (n=23, lower panel) or third (n=79, upper panel) trimester, delivery, and postpartum, with samples submitted for WGS and/or 16S rRNA targeted amplicon analysis. An additional 80 gravidae were enrolled and sampled only at delivery, with samples submitted for 16S rRNA targeted amplicon analysis. 83 of the 102 longitudinally enrolled gravidae had samples submitted for WGS analysis. (**B**) Per subject details regarding time point, sampling subsite acquisition, and method of sequencing for subjects enrolled longitudinally and submitted and analyzed for WGS and/or targeted 16S rRNA gene amplicon analysis.

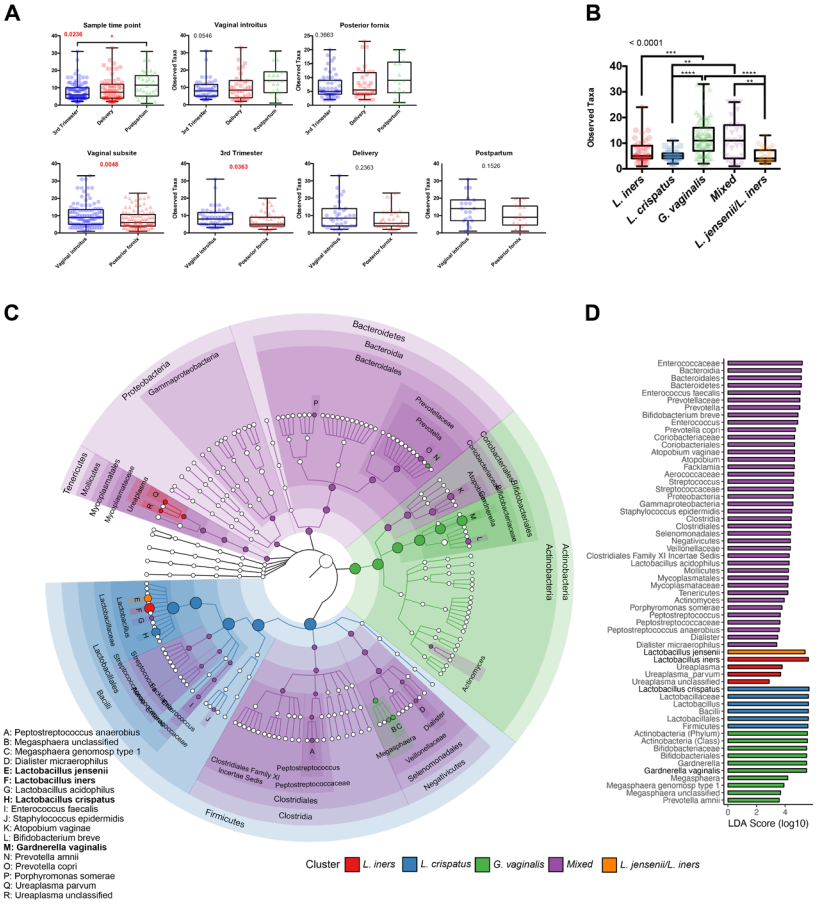



**Figure S2.** Vaginal microbiome alpha diversity and taxonomic associations with cluster membership, Related to Figure 1. (**A**) Observed species across perinatal sample time points and vaginal subsites. Samples were stratified by sample time point, and further subset by vaginal subsite. For samples stratified by sample time point a Kruskal-Wallis test (p value in upper left quadrant of the plot) was first performed, followed by Dunn's post-test ($p \leq 0.05$ denoted by the asterisk). For samples stratified by vaginal subsite a Mann-Whitney test was performed. All significant p values ($p \leq 0.05$) are highlighted in red, with non-significant in black. (**B**) Observed species within predominant taxa clusters. Kruskal-Wallis test was first performed (p value in upper left quadrant of the plot), followed by Dunn's post-test (significance denoted by asterisk: * =$p \leq 0.05$, ** =$p \leq 0.01$, *** =$p \leq 0.001$, and **** =$p \leq 0.0001$). (**C-D**) Taxonomic associations with cluster membership. The effect size for differentially enriched taxa based on k-means cluster assignment of samples was calculated using LEfSe. (**C**) Cladogram of taxa enriched within the five k-means clusters. The *L. iners*, *L. crispatus*, *L. jensenii/L. iners*, and *G. vaginalis* clusters were enriched for their respective predominant taxa, whereas the mixed cluster was enriched for BV associated taxa. Node size reflects the relative abundance of taxa and shading reflects cluster membership. (**D**) LDA scores for enriched taxa, colored by cluster membership. Keystone species are bolded.

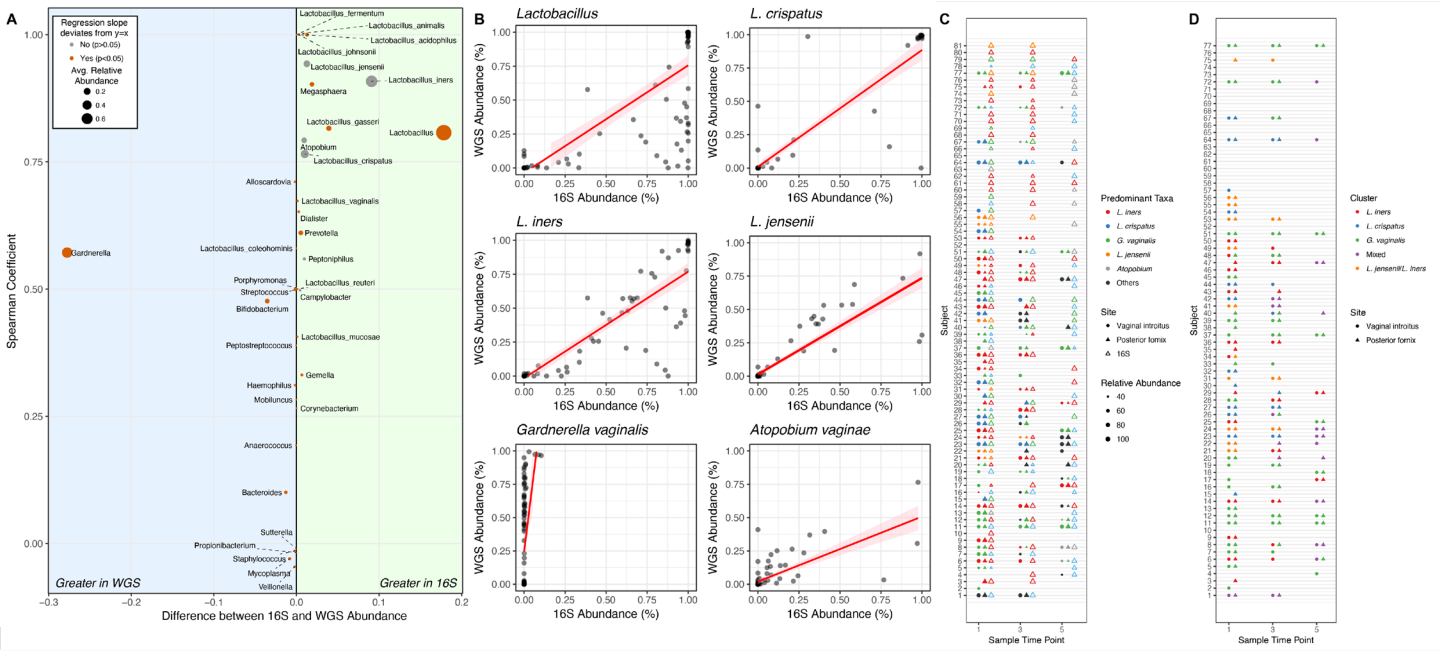



**Figure S3.** Comparison of WGS metagenomics and 16S (V1V3 hypervariable region) taxonomic relative abundances, Related to Figure 1. (**A**) Comparisons of the relative abundances for each identified taxa as determined by whole genome shotgun sequencing (WGS) and 16S V1V3 sequencing. X-axis indicates the mean difference in relative abundance between WGS and 16S datasets for each indicated species (see key, A) while Y-axis indicates the spearman coefficient between the datasets. The size of the circle in panel A is proportional to the average relative abundance. Taxa whose correlation significantly differed from a y=x model ($p<0.05$, indicating perfect 1:1 correlation between the WGS and 16S datasets) are colored in red. (**B**) Scatter plots for *Lactobacillus* species, *Gardnerella* and *Atopobium* are shown comparing the relative abundance of each taxa as determined by WGS (y-axis) and 16S (x-axis) within the same samples. Linear regression lines are provided with a 95% confidence interval. (**C-D**) Per subject comparison of the predominant taxa called by WGS metagenomics versus 16S (V1V3 hypervariable region). (**C**) WGS/16S samples highlighted by predominant taxa. (**D**) WGS samples clustered by k-means. Sample time points are denoted as 1 – third trimester, 3 – delivery, and 5 – postpartum.

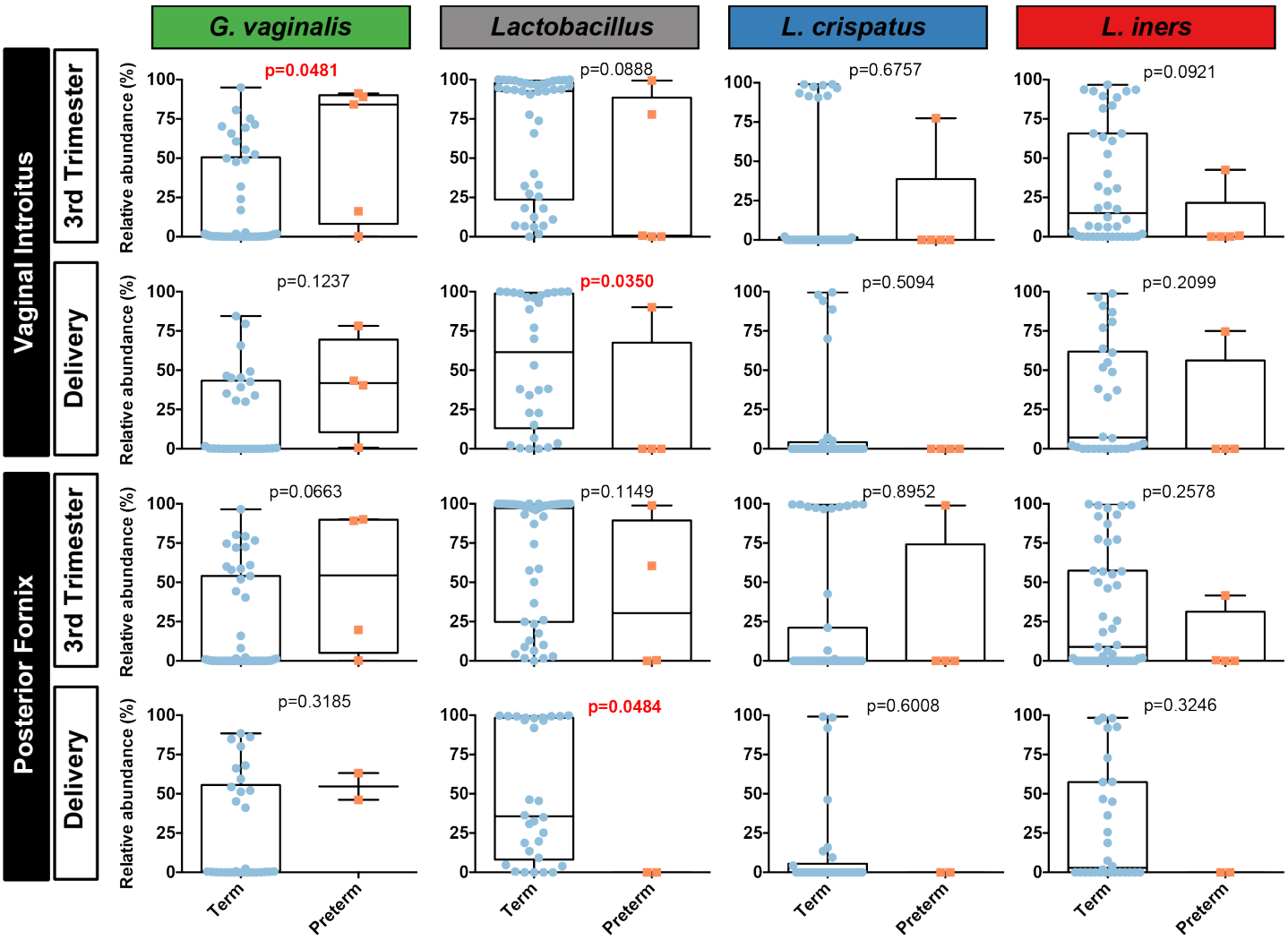



**Figure S4.** WGS relative abundance of *G. vaginalis*, *Lactobacillus* genera, *L. crispatus*, and *L. iners* stratified by vaginal subsite and sampling time during pregnancy, Related to Figure 4. Significant p values ($p \leq 0.05$) derived from the non-parametric Mann-Whitney test are highlighted in red, non-significant are in black.

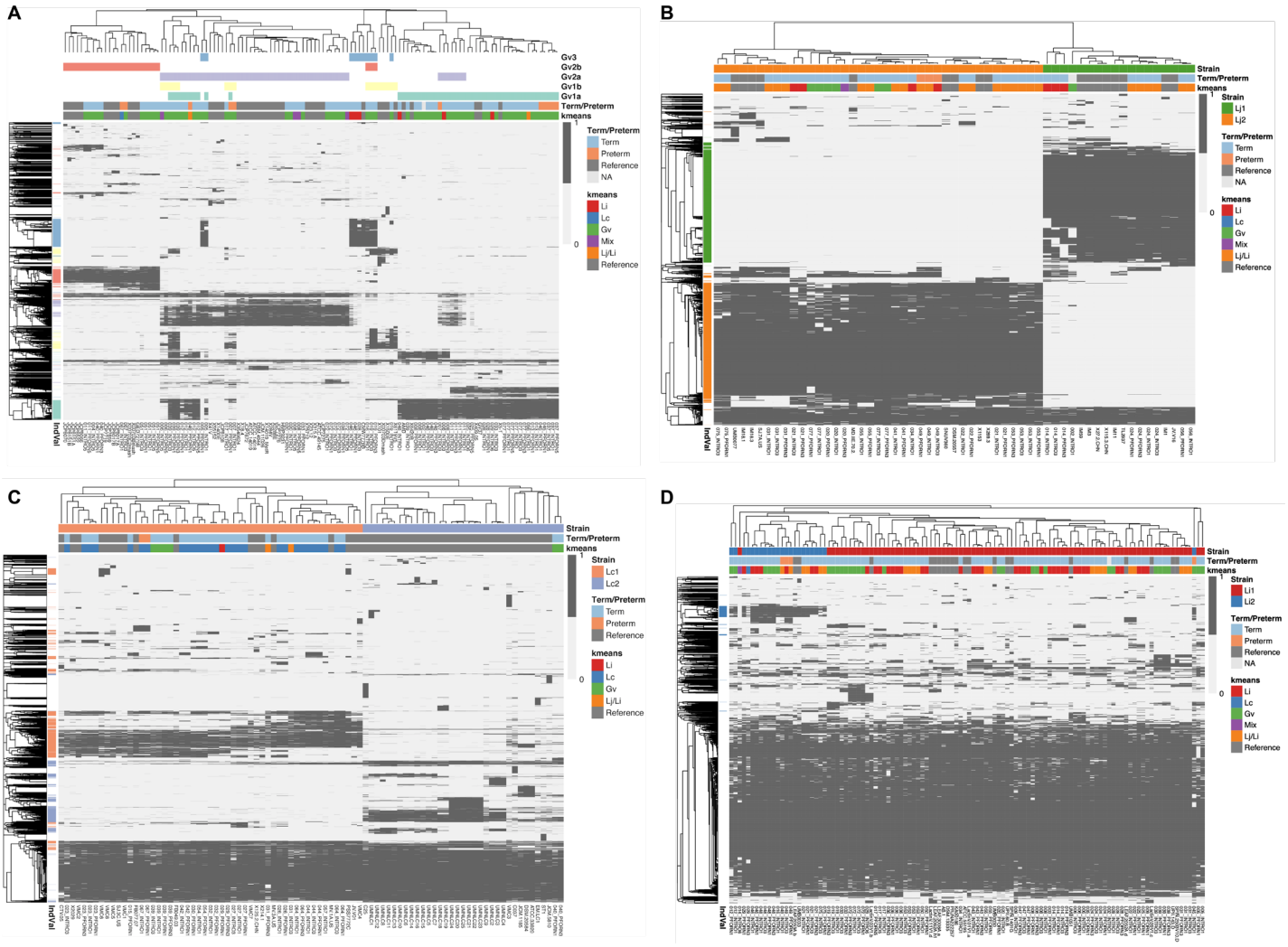



**Figure S5.** Strain level pangenome centroid heatmaps for vaginal samples from the perinatal interval, Related to Figure 6. (**A**) *Gardnerella vaginalis*; (**B**) *Lactobacillus crispatus*; (**C**) *Lactobacillus iners*; and (**D**) *Lactobacillus jensenii*. Binary Jaccard clustering for reference genomes and samples was performed on rows and columns. Rows annotations indicate IndVal determined strain-specific pangenome centroids (presence = 1, absence = 0). Columns are annotated with strain assignments, term/preterm birth outcomes, and kmeans cluster membership. Sample and reference genomes are indicated at the bottom; sample name format is as follows: subject ID, vaginal subsite, and time of collection, with the following abbreviations used for vaginal subsite and time of collection: INTRO – vaginal introitus, PFORN – posterior fornix, 1 – third trimester, 3 – delivery, 5 – postpartum.

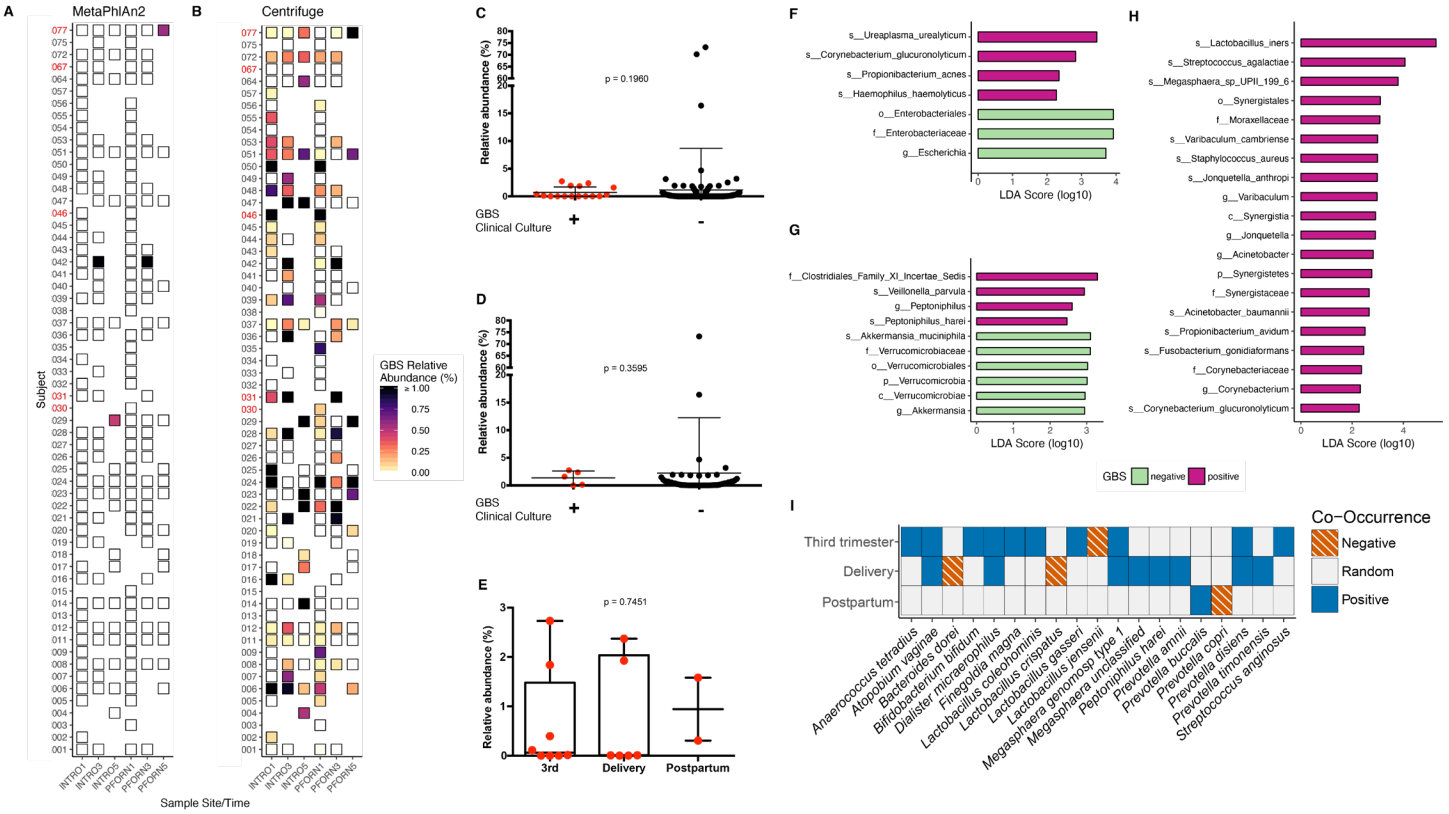



**Figure S6.** Group B *Streptococcus* in the perinatal interval, Related to Figure 7. (**A-B**) Comparison of MetaPhlAn2 and Centrifuge assignments of GBS relative abundance across 201 vaginal samples. (**A**) MetaPhlAn2 classified four samples with a greater than zero relative abundance of GBS. (**B**) Centrifuge classified 123 samples with a greater than zero relative abundance of GBS. Subjects are aligned across the y-axis, with GBS subjects that had a positive clinical culture highlighted in red. (**C**) Relative abundance of all samples, (**D**) maximal relative abundance per subject, and (**E**) relative abundance of samples over time from subjects with positive GBS clinical cultures. Mann-Whitney test p values are given for C and D, and the Kruskal-Wallis p value is given for C. Vaginal sample sites and times are aligned across the x-axis and denoted as INTRO – vaginal introitus, PFORN – posterior fornix, 1 – third trimester, 3 – delivery, 5 – postpartum. (**F-H**) LEfSe results on differential enrichment of taxa during pregnancy based on vaginal GBS status by clinical culture status (rectovaginal swab) or metagenomic classification. Average gestational relative abundance values were used for LEfSe. (**F**) Subjects were classified according to GBS clinical culture status. (**G**) Subjects were classified as GBS positive based on a metagenomic relative abundance greater than zero. (**H**) Subjects were classified as GBS positive based on a metagenomic relative abundance greater than one percent. **(I)** Significant pairwise co-occurrences for vaginal GBS at the third trimester, delivery, and postpartum perinatal time points.

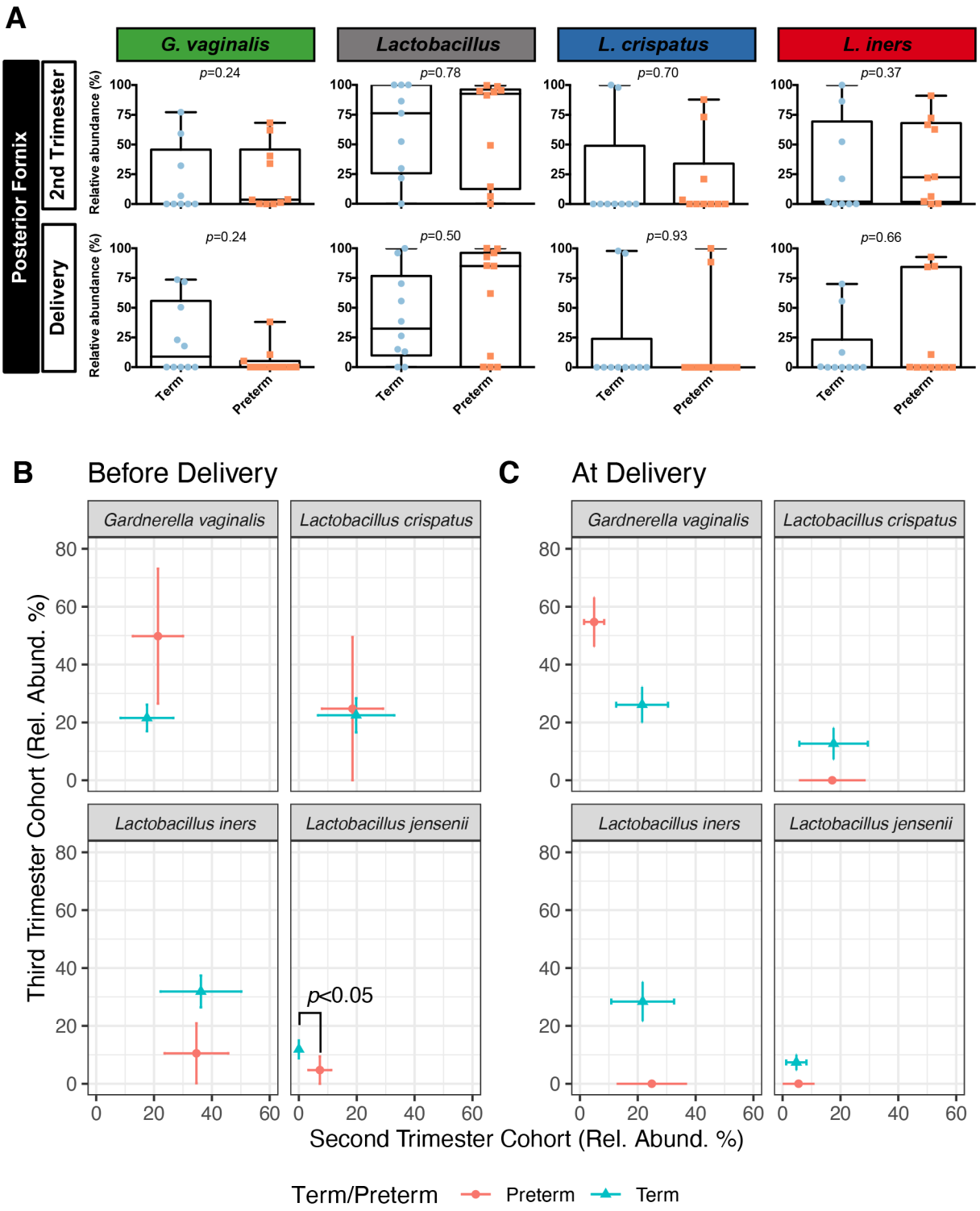



**Figure S7.** Comparisons of the relative abundance of vaginal *G. vaginalis*, *Lactobacillus* genera, *L. crispatus*, and *L. iners*, Related to STAR Methods. (**A**) Abundance of keystone species in the second trimester cohort stratified by pregnancy and delivery sampling timepoints. **(B-C)** Comparison of the relative abundances of *G. vaginalis*, *L. crispatus*, *L. iners*, and *L. jensenii* present in the posterior fornix by cohort. (**B**) Relative abundance of keystone species during pregnancy. (**C**) Relative abundance of keystone species at delivery. P values are derived from the non-parametric Mann-Whitney test.

## Supplemental Tables

| | Preterm (n) | Term (n) | Fisher's exact test, p | Odds ratio (95% CI) |
|---|---|---|---|---|
| **Gv1** | 2/3 | 4/10 | 0.56 | 3.0 (0.2 – 45.3) |
| **Gv1a** | 2/3 | 4/10 | 0.56 | 3.0 (0.2 – 45.3) |
| **Gv1b** | 0/3 | 0/10 | NA | NA |
| **Gv2** | 1/3 | 5/10 | >0.99 | 0.7 (0.1 – 10.02) |
| **Gv2a** | 0/3 | 3/10 | 0.53 | 0.3 (0.01 – 7.7) |
| **Gv2b** | 1/3 | 2/10 | >0.99 | 2.0 (0.1 – 34.9) |
| **Gv3** | 0/3 | 1/10 | >0.99 | 0.9 (0.03 – 27.9) |
| **Lc1** | 1/1 | 14/15 | >0.99 | 0.3 (0.01 – 11.5) |
| **Lc2** | 0/1 | 1/15 | >0.99 | 3.2 (0.1 – 119.8) |
| **Li1** | 0/2 | 6/32 | >0.99 | 0.8 (0.03 – 19.1) |
| **Li2** | 2/2 | 26/32 | >0.99 | 1.2 (0.1 – 28.8) |
| **Lj1** | 0/1 | 3/14 | >0.99 | 1.1 (0.04 – 33.4) |
| **Lj2** | 1/1 | 11/14 | >0.99 | 0.1 (0.003 – 3.1) |

**Table S2.** Frequency of association and odds ratios for keystone species strains, present alone, during pregnancy and preterm birth outcomes, Related to Figure 6.

| | Preterm (n) | Term (n) | Fisher's exact test, p | Odds ratio (95% CI) |
|---|---|---|---|---|
| **Gv1** | 3/4 | 2/10 | 0.10 | 12.0 (0.8 – 186.5) |
| **Gv1a** | 2/4 | 1/10 | 0.18 | 9.0 (0.5 – 155.4) |
| **Gv1b** | 2/4 | 1/10 | 0.18 | 9.0 (0.5 – 155.4) |
| **Gv2** | 1/4 | 2/10 | >0.99 | 1.3 (0.1 – 20.7) |
| **Gv2a** | 1/4 | 1/10 | 0.51 | 3.0 (0.1 – 64.3) |
| **Gv2b** | 0/4 | 1/10 | >0.99 | 0.7 (0.02 – 20.9) |
| **Gv3** | 0/4 | 1/10 | >0.99 | 0.7 (0.02 – 20.9) |
| **Lc1** | 3/9 | 2/10 | 0.63 | 2.0 (0.3 – 16.0) |
| **Lc2** | 1/9 | 0/10 | >0.99 | 3.3 (0.1 – 91.7) |
| **Li1** | 5/8 | 5/9 | >0.99 | 1.3 (0.2 – 9.3) |
| **Li2** | 2/8 | 1/9 | 0.58 | 2.7 (0.2 – 36.8) |
| **Lj1** | 3/9 | 0/10 | 0.09 | 11.31 (0.5 – 256.4) |
| **Lj2** | 1/9 | 0/10 | 0.47 | 3.7 (0.1 – 103.2) |

**Table S3.** Frequency of association and odds ratios for keystone species strains, present alone, during pregnancy and preterm birth outcomes in the Second trimester validation cohort – during mid-pregnancy timepoint, Related to Figure 6.

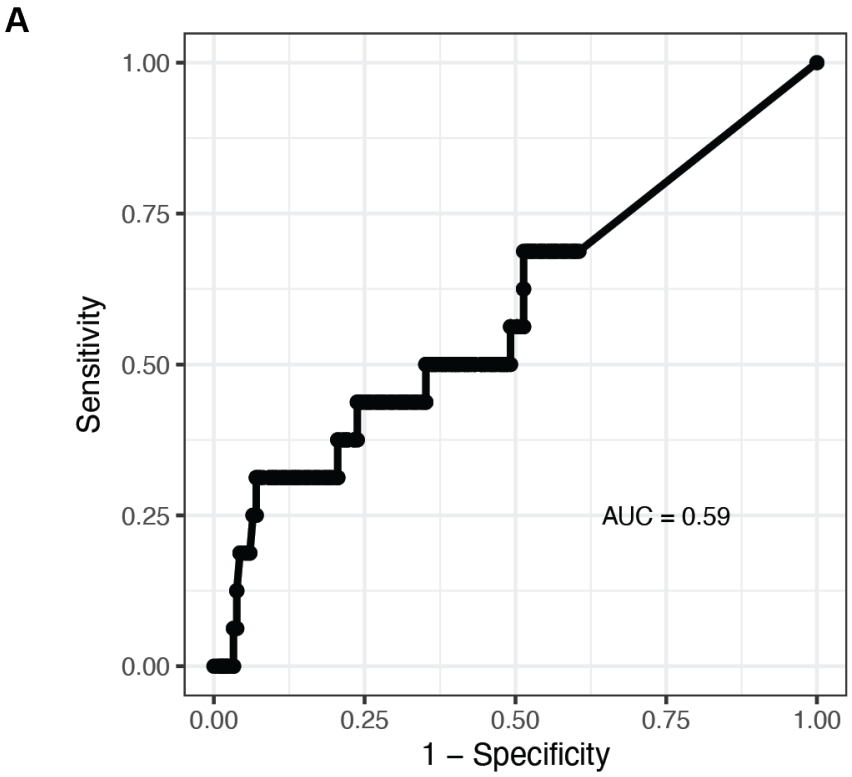
| | Participants with WGS Sequencing | | | | Participants with 16S Sequencing | | |
|---|---|---|---|---|---|---|---|
| | Term (n=55) | Preterm (n=5) | p | | Term (n=142) | Preterm (n=17) | p |
| **Maternal Characteristics** | | | | | | | |
| Age | 30.3±5.0 | 22.8±5.3 | 0.002 | | 30.5±5.6 | 25.3±6.7 | <0.001 |
| Ethnicity* | | | | | | | |
| *Hispanic* | 46 | 5 | 0.65 | | 122 | 14 | 0.9 |
| *African-American* | 5 | 0 | | | 11 | 2 | |
| *Asian* | 3 | 0 | | | 7 | 1 | |
| *Caucasian* | 0 | 0 | | | 2 | 0 | |
| BMI | 28.9±6.1 | 28.0±6.0 | 0.76 | | 27.8±6.0 | 26.9±6.2 | 0.59 |
| Gravidity | 3.4±1.7 | 3.0±2.3 | 0.6 | | 2.2±1.6 | 2.9±1.9 | 0.6 |
| Parity | 1.9±1.5 | 1.4±1.7 | 0.49 | | 1.7±1.4 | 1.5±1.5 | 0.58 |
| **Antepartum Clinical Diagnoses of GBS positive culture, BV, or Gestational diabetes** | | | | | | | |
| GBS culture positive | 4 (7.4%) | 1 (20%) | 0.33 | | 14 (9.8%) | 3 (17.6%) | 0.33 |
| Bacterial Vaginosis | 5 (9.3%) | 2 (40%) | 0.04 | | 10 (7.0%) | 4 (23.5%) | 0.02 |
| Gestational Diabetes | 26 (48.1%) | 0 (0%) | 0.03 | | 45 (31.7%) | 1 (5.8%) | 0.03 |
| **Neonatal Characteristics** | | | | | | | |
| Cesarean Born | 17 (31.4%) | 3 (60%) | 0.15 | | 48 (33.2%) | 6 (35.2%) | 0.9 |
| Gestational Age at delivery (Weeks) | 38.8±1.1 | 33.8±2.7 | <0.001 | | 38.8±1.6 | 34.0±3.5 | <0.001 |
| Sex | | | | | | | |
| *Male* | 28 | 1 | 0.23 | | 73 | 11 | 0.32 |
| *Female* | 26 | 4 | | | 69 | 9 | |

**Table S4.** Third trimester cohort demographics. *, missing data for 1 participant, Related to STAR Methods.

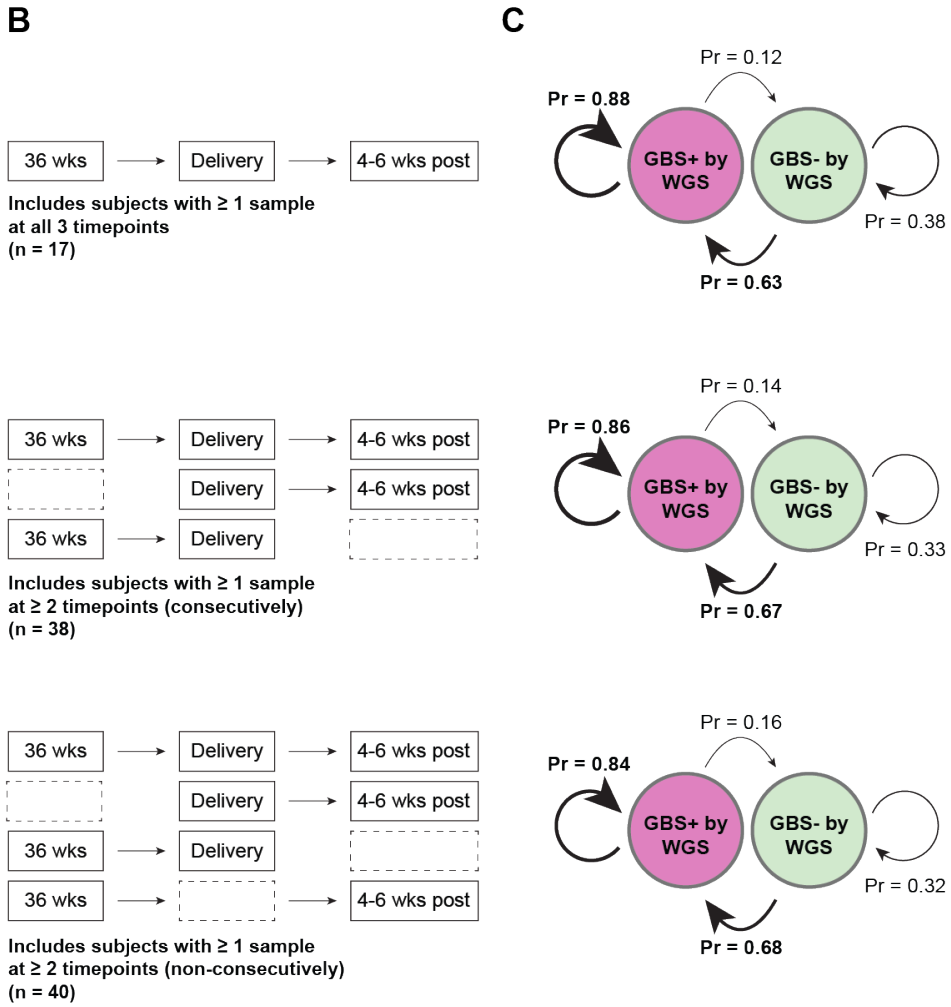
| | Participants with WGS Sequencing | | |
|---|---|---|---|
| | Term (n=12) | Preterm (n=11) | p |
| **Maternal Characteristics** | | | |
| Age | 30.2±4.5 | 33.45±5.3 | 0.15 |
| Ethnicity | | | |
| *Hispanic* | 8 | 9 | 0.69 |
| *African-American* | 2 | 2 | |
| *Asian* | 0 | 0 | |
| *Caucasian* | 2 | 0 | |
| BMI | 28.4±7.2 | 31.7±8.6 | 0.34 |
| Gravidity | 4.0±1.4 | 5.1±3.8 | 0.47 |
| Parity | 2.4±1.1 | 3.4±3.5 | 0.58 |
| **Antepartum Clinical Diagnoses of GBS positive culture, BV, or Gestational diabetes** | | | |
| GBS | 1 (9.1%)[1] | 1 (11.1%)[2] | 0.88 |
| Bacterial Vaginosis | 1 (8.3%) | 4 (36.4%) | 0.10 |
| Gestational Diabetes | 3 (25%) | 2 (18.1%) | 0.69 |
| **Neonatal Characteristics** | | | |
| Cesarean Delivery | 3 (25%) | 3 (27.3%) | 0.90 |
| Gestational Age (Weeks) | 38.1±1.3 | 33.3±3.5 | <0.0001 |
| Sex | | | |
| *Male* | 8 | 4 | 0.15 |
| *Female* | 4 | 7 | |

**Table S5.** Second trimester cohort demographics, Related to STAR Methods.
[1]testing not done for n=1; [2]testing not done for n=2.

## Supplemental Methods




| | SENS | SPEC | PPV | NPV |
|---|---|---|---|---|
| **WGS GBS+, 1 time point** | 0.80 | 0.11 | 0.08 | 0.86 |
| **WGS GBS+, 2 time points** | 0.67 | 0.46 | 0.09 | 0.94 |
| **WGS GBS+, 3 time points** | 1.0 | 0.47 | 0.11 | 1.0 |




**Methods S1.** Metagenomics as a reliable screening tool for GBS status as compared in Figure 7. (A) ROC curve analysis for the accuracy of metagenomic classification of GBS for all samples based on clinical classification. Sensitivity (SENS), specificity (SPEC), positive predictive value (PPV), and negative predictive value (NPV) were calculated based on the number of time points WGS samples were collected and determined as GBS positive by WGS metagenomics. (B-C) Transition probabilities for vaginal GBS status in the perinatal period. (B) Strategy for calculating transition probabilities. (C) Transition probabilities for vaginal GBS status by WGS metagenomics.