Title: Evidence that vertical transmission of the vaginal microbiota can persist into adolescence

Running title: Vertical transmission of the vaginal microbiota

Authors: Christine M. Bassis (cbassis@umich.edu)[1], Kaylie A. Bullock (bullockk@umich.edu)[1], Daniel E. Sack (sackd@umich.edu)[1], Katie Saund (katiephd@umich.edu)[2], Ali Pirani (apirani@med.umich.edu)[2], Evan S. Snitkin (esnitkin@med.umich.edu)[1,2], Veronica I. Alaniz (veronica.alaniz@ucdenver.edu)[3], Elisabeth H. Quint (equint@med.umich.edu)[3], Vincent B. Young (youngvi@med.umich.edu)[1,2] and Jason D. Bell (jasonbel@med.umich.edu)[3].

1. Division of Infectious Disease, Department of Internal Medicine
2. Department of Microbiology and Immunology
3. Department of Obstetrics and Gynecology, University of Michigan, Ann Arbor, MI 48109

Corresponding author: Christine M. Bassis (cbassis@umich.edu)

Abstract

Background: Factors that influence vaginal microbiota composition, including its source, are not well understood. To determine if vaginal microbiota transmission from mother to daughter at birth influences the human vaginal microbiota composition in adolescence, we investigated the relationship between the vaginal microbiota of 13 mother/daughter pairs and the daughter's birth mode.

Results: Based on analysis of bacterial 16S rRNA gene sequences, the vaginal microbiotas of mother/daughter pairs were more similar to each other if the daughter was born by vaginal delivery rather than by C-section. Additionally, genome sequences from an important member of the vaginal microbiota, *Lactobacillus crispatus*, isolated from one mother/daughter pair in which the daughter was born by vaginal delivery, were highly similar.

Conclusions: Both community-level analysis and isolate genome sequence analysis provide evidence of birth-mode dependent transmission and persistence of at least some members of the vaginal microbiota.

Importance

The composition of the human vaginal microbiota is related to many aspects of health from infection susceptibility to preterm birth. Our study provides evidence that transmission of vaginal bacteria from mother to daughter at birth is an important factor influencing vaginal microbiota composition into adolescence.

Keywords

Vaginal microbiota, transmission, birth mode, 16S rRNA gene sequences, *Lactobacillus crispatus* genomics

47 **Background**

48 The vaginal microbiota plays an important role in human health. The community
49 structure of the vaginal microbiota has strong links to HIV susceptibility and preterm
50 birth [1-3]. While there has been considerable effort in examining the role of the
51 gastrointestinal microbiota in human health and disease (e.g. [4]), there have been
52 fewer studies examining the microbial communities that inhabit the female reproductive
53 tract. The composition of the vaginal microbiota is distinct from other body sites and
54 contains types of bacteria that seem specific to the vagina [5]. For example, the vaginal
55 microbiota is often dominated by specific types of *Lactobacillus,* most commonly *L.*
56 *crispatus* and *L. iners* [6, 7]. Vaginal *Lactobacillus* sp. are thought to maintain
57 dominance and inhibit colonization of other microbes through lactic acid production [8,
58 9].

59 Despite strong evidence that the vaginal microbiota can have significant impacts
60 on a woman's reproductive tract health, the factors that influence the composition of the
61 vaginal microbiota are not well understood. It is not known how this vagina-specific
62 community is maintained from generation to generation. One possibility is that at least
63 some members of the vaginal microbiota are transmitted from mother to daughter at
64 birth and maintained in daughters through adolescence.

65 In healthy babies, the first large, direct exposure to microbes occurs at birth. Birth
66 mode has been shown to influence the composition of the newborn microbiota (gut,
67 skin, mouth), likely due to different bacterial exposure in vaginal delivery and C-sections
68 [10, 11]. However, the effect of birth mode on the composition of the vaginal microbiota
69 has not been investigated. In this study, we compared the vaginal microbiotas of 13

70 mother/daughter pairs and investigated the effect of birth mode on mother/daughter
71 microbiota similarity. We also compared the genome sequences from *Lactobacillus*
72 *crispatus* isolates from one mother/daughter pair. We hypothesized that the vaginal
73 microbiota of mothers and daughters would be more similar if the daughter was born by
74 vaginal delivery than by C-section.

75 **Methods**

76 **Subject recruitment and sample collection**

77 Mother/daughter pairs were recruited from the Pediatric and Adolescent
78 Gynecology Clinic at the University of Michigan Health System in 2014 and 2015.
79 Exclusions were pregnancy and age of less than 15 years. Written, informed consent
80 was obtained and participants completed a baseline survey on their demographics and
81 pertinent gynecologic and medical history. Vaginal samples were self-collected using a
82 dual-headed swab (Starplex Scientific, S09D) at baseline and then weekly for 4 weeks.
83 The baseline swab was obtained in the clinic, with immediate storage on ice and
84 transfer to -80°C within a few hours. The subsequent swabs were returned via mail at
85 ambient temperature. After the fifth swab was received and a completion incentive was
86 mailed to the subject, the link between samples and subject names was destroyed,
87 irreversibly de-identifying all samples. The study was approved by the University of
88 Michigan IRB (HUM00086661).

89 **DNA isolation and 16S rRNA gene sequencing**

90 One of the swab heads from each sample was clipped directly into the bead plate
91 of a PowerMag Microbiome RNA/DNA Isolation Kit (Mo Bio Laboratories, Inc.). DNA
92 isolation was performed according the manufacturer's instructions using an epMotion

93 5075 liquid handling system. The V4 region of the 16S rRNA gene was amplified from 1
94 or 7μl DNA and sequenced with a MiSeq (Illumina, San Diego, CA) using the 500 cycle
95 MiSeq Reagent Kit, v. 2 (Illumina, catalog No. MS-102–2003) by the University of
96 Michigan Microbial Systems Molecular Biology Laboratory as described previously
97 [12]. The other swab head was used for cultivation or stored at -80°C.

98 **Bacterial community analysis**

99 The 16S rRNA gene sequences were processed using mothur v.1.36.1 and
100 v.1.39.5 following the mothur MiSeq SOP [13, 14]. Details of the processing steps are
101 available in mother.daughter_mothur.batch
102 (https://github.com/cbassis/MotherDaughter_Vaginal_Microbiota.study). After sequence
103 processing and alignment to the SILVA reference alignment (Release 102) [15],
104 sequences were binned into operational taxonomic units (OTUs) based on 97%
105 sequence similarity using the average neighbor method [16, 17]. Samples with fewer
106 than 1000 sequences were excluded from the analysis. OTUs were classified to the
107 genus level within mothur using a modified version of the Ribosomal Database Project
108 (RDP) training set (version 9) [18, 19]. To further classify the *Lactobacillus* OTUs,
109 representative sequences were analyzed using standard nucleotide BLAST for highly
110 similar sequences (megablast) on the National Center for Biotechnology Information
111 (NCBI) BLAST web page (https://blast.ncbi.nlm.nih.gov/Blast.cgi) [20]. OTU relative
112 abundances were calculated and plotted in a heatmap. To compare bacterial
113 communities between pairs, within pairs and within subjects, we calculated θ_{YC}
114 distances (a metric that takes relative abundances of both shared and non-shared
115 OTUs into account) [21]. A Kruskal-Wallis test with a Dunn's posttest or a Wilcoxon

116 (Mann-Whitney) test were used to determine if differences in θ_{YC} distances were
117 statistically significant. Principal coordinates analysis (PCoA) was used to visualize the
118 θ_{YC} distances between samples. R Studio (Version 1.1.456) with R (Version 3.5.1) was
119 used for the statistical tests and plotting the heat map, box and whisker plots, and the
120 ordination using the code available:
121 https://github.com/cbassis/MotherDaughter_Vaginal_Microbiota.study/tree/master/R_co
122 de. Adobe Illustrator (CS6) was used for labeling and formatting figures.

123 ***Lactobacillus crispatus* isolation**

124 For pair I, the second swab head from the freshly collected baseline vaginal
125 sample was swabbed onto an MRS agar plate and incubated in an anaerobic chamber
126 (Coy Laboratory Products) at 37°C. Individual isolates were identified via Sanger
127 sequencing of the near-full length 16S rRNA gene.

128 **DNA isolation and genome sequencing**

129 Three *Lactobacillus crispatus* isolates from pair I, 2 from the mother and 1 from
130 the daughter, were grown overnight in 1 ml liquid MRS in an anaerobic chamber (Coy
131 Laboratory Products) at 37°C. Genomic DNA was isolated from the liquid cultures using
132 the PowerMicrobiome™ RNA Isolation Kit (Mo Bio Laboratories, Inc.) without the DNase
133 treatment. Genome sequencing was performed by the Microbial Systems Molecular
134 Biology Laboratory at the University of Michigan using an Illumina Nextera™ sequencing
135 kit and a MiSeq (Illumina, San Diego, CA).

136 **Genome sequence analysis**

137 Phylogenetic relationships between *L. crispatus* isolates from mother/daughter pair I
138 and all *L. crispatus* strains with genome sequences available as fastq files from NCBI

139 on December 27th, 2018 were determined based on recombination-filtered single
140 nucleotide polymorphisms (SNPs). Quality of reads was assessed with FastQC v0.11.3
141 [22], and Trimmomatic 0.36 [23] was used for trimming adapter sequences and low-
142 quality bases. Variants were identified by (i) mapping filtered reads to reference
143 genome sequence *L. crispatus* ST1 (SAMEA2272191) using the Burrows-Wheeler
144 short-read aligner (bwa-0.7.17) [24, 25], (ii) discarding polymerase chain reaction
145 duplicates with Picard (picard-tools-2.5.0) [26], and (iii) calling variants with SAMtools
146 (samtools-1.2) and bcftools [27]. Variants were filtered from raw results using GATK 's
147 (GenomeAnalysisTK-3.3-0) VariantFiltration (QUAL, >100; MQ, >50; >=10 reads
148 supporting variant; and FQ, <0.025) [28]. In addition, a custom python script was used
149 to filter out single-nucleotide variants that were (i) <5 base pairs (bp) in proximity to
150 indels, (ii) fell under Phage and Repeat region of the reference genome (identified using
151 Phaster [29] and Nucmer (MUMmer3.23) [30]), (iii) not present in the core genome, or
152 (iv) in a recombinant region identified by Gubbins 2.3.1 [31]. A maximum likelihood tree
153 was constructed in RAxML 8.2.8 [32] using a general-time reversible model of sequence
154 evolution. Bootstrap analysis was performed with the number of bootstrap replicates
155 determined using the bootstrap convergence test and the autoMRE convergence
156 criteria (-N autoMRE). Bootstrap support values were overlaid on the best scoring tree
157 identified during rapid bootstrap analysis (-f a). The final maximum likelihood tree was
158 plotted and pairwise SNP distances were calculated in R Studio (Version 1.1.463) with
159 R (Version 3.5.3):
160 https://github.com/cbassis/MotherDaughter_Vaginal_Microbiota.study/blob/master/R_co

161 de/Mother_Daughter_Figure_3_Genome_Tree_and_Genome_Analysis.Rmd. Adobe
162 Illustrator (CS6) was used for labeling and formatting the figure.

163 **Calculation of doubling time estimate for vaginal *L. crispatus in vivo***

164 We used the number of SNPs between the pair I mother and daughter *L.*
165 *crispatus* isolates to estimate the doubling time of vaginal *L. crispatus in vivo* if all SNPs
166 in the recombination-filtered core genome were due to mutations acquired since the
167 daughter's birth:

168 Doubling time=(mutation rate)(daughter's age)(genome length)/(# of mutations)

169 The mutation rate of *L. crispatus* is unknown, so for this estimate we used the published
170 mutation rate of another *Lactobacillus, L. casei* Zhang, *in vitro,* without antibiotics
171 (1.0×10^{-9} bp/generation) [33]. The pair I daughter's age in hours was: 175,200 hours
172 =(20 years)(365 days/year)(24 hour/day). The average length of the recombination-
173 filtered core genome (940,943 bp) was used for genome length. We assumed that the
174 isolates arose from a common ancestor and that all mutations were non-convergent, so
175 the number of mutations acquired would equal the number of SNPs between the
176 mother's isolate and the daughter's isolate divided by 2.

177 **Results**

178 **Subject characteristics and sequencing results**

179 A total of 107 self-collected, vaginal swab samples were obtained from 26
180 subjects (13 mother/daughter pairs) (Table 1). Each subject returned 1-5 weekly
181 samples (median=5 samples/subject, IQR=1). After sequence processing and exclusion
182 of samples with fewer than 1000 sequences, a total of 2,336,437 high quality bacterial
183 16S rRNA gene sequences from 101 samples were analyzed with an average of 23,133

184 +/- 10,212 sequences per sample.

Table 1. Subject Characteristics

	Mother (n=13)	Daughter (n=13)
	n (%)	n (%)
Age, mean ± SD, years	47±6	17±2
Race: White (vs. Black, Asian, Hispanic, other)	12 (92)	12 (92)
Birth mode: Vaginal (vs. C-section)	10 (77)	10 (77)
Reproductive stage: Premenarchal	0 (0)	1 (8)
Reproductive stage: Reproductive	8 (62)	12 (92)
Reproductive stage: Postmenopausal	5 (38)	0 (0)

185

186 **An individual's vaginal microbiota is relatively stable over 4 weeks**

187 During the sampling period, the vaginal microbiota of each subject was relatively
188 stable. The high stability of the vaginal microbiota is apparent from the consistent within
189 subject community composition (Figure 1). For example, high relative abundances of
190 OTU1 (*L. crispatus*) and/or OTU2 (*L. iners*) persisted from week to week in many
191 subjects. Additionally, average θ_{YC} distances were significantly lower within subjects
192 than between subjects (Figure 2A) and samples clustered by subject in a PCoA based
193 on θ_{YC} distances (Supplemental Figure 1).

194 **Daughters born via vaginal delivery have greater microbiota similarity with their**
195 **mothers than those born via C-section**

196 To determine if mothers and their daughters had more similar vaginal microbiotas
197 than unrelated subjects, we compared the average θ_{YC} distances between all unrelated
198 subjects (between pairs) and the average θ_{YC} distances between mothers and their own
199 daughters (within pairs) (Figure 2A). There was a trend toward greater similarity (lower
200 θ_{YC} distances) within all mother/daughter pairs than between subjects in different

201 mother/daughter pairs. To determine if birth mode was related to vaginal microbiota
202 similarity within mother/daughter pairs, we compared the average within pair θ_{YC}
203 distances for pairs in which the daughter was born by vaginal delivery and by C-section
204 (Figure 2B). The average within pair θ_{YC} distances were significantly lower for pairs in
205 which the daughter was born by vaginal delivery compared to C-section (Fig. 2B).
206 Therefore, the vaginal microbiotas of daughters born by vaginal delivery were
207 significantly more similar to their mothers' than the daughters born by C-section were to
208 their mothers' (Fig. 2B).

209 ***Lactobacillus crispatus* isolates from mother/daughter pair I have highly similar**
210 **genome sequences**

211 The birth mode-dependent similarity of the vaginal microbiotas of mothers and
212 their daughters suggested that vaginal bacteria could be transmitted between
213 generations at birth and persist into adolescence. However, it is possible that genetic or
214 environmental factors shared by a mother and her daughter lead to acquisition of similar
215 bacteria later, resulting in the *de novo* establishment of similar vaginal communities. To
216 investigate the possibility of direct transmission and persistence of one member of the
217 vaginal microbiota, we generated draft genome sequences of *Lactobacillus crispatus*
218 strains isolated from the freshly collected second swab head of mother/daughter pair I.
219 The draft genome sequences of these isolates were compared with publicly available *L.*
220 *crispatus* genome sequences by constructing a maximum likelihood phylogenetic tree
221 based on a recombination-filtered core genome alignment. Interestingly, the three
222 strains of *L. crispatus* from mother/daughter pair I, UMP1M1, UMP1M2 and UMP1D1,
223 were more similar to each other than to any of the other strains, including others

224 isolated from the female reproductive tract (Fig.3).

225 We also calculated the number of SNPs between our isolates using the
226 recombination-filtered core genome alignment. There were 11 recombination-filtered
227 SNPs between the 2 isolates from the mother (UMP1M1 and UMP1M2) and 25 and 16
228 recombination-filtered SNPs between the daughter's isolate (UMP1D1) and the 2 isolates
229 from the mother (UMP1M1 and UMP1M2, respectively).

230 **Estimate of *in vivo* doubling time for vaginal *L. crispatus***

231 To further investigate the plausibility that the *L. crispatus* strain isolated from
232 daughter I descended from a strain transmitted from her mother at birth, we estimated
233 the doubling time that would allow our isolates to acquire the observed number of SNPs
234 over 20 years. Based on the 25 SNPs between UMP1M1 and UMP1D1, the estimated
235 doubling time for *L. crispatus in vivo* would be 13.2 hours. Based on the 16 SNPs
236 between UMP1M2 and UMP1D1, the estimated doubling time would be 20.6 hours.

237 **Discussion**

238 Our study provides preliminary evidence that the vaginal microbiota is vertically
239 transmitted from mother to daughter at birth via vaginal delivery and persists into
240 adolescence. Because the daughters in our study were 15-21 years old, both
241 transmission and persistence were required to observe evidence of vertical
242 transmission. The first piece of evidence supporting vertical transmission is that the
243 vaginal microbiotas of mothers and their adolescent daughters were more similar if their
244 daughter was born by vaginal delivery rather than C-section. The second piece of
245 evidence supporting vertical transmission and persistence is that an important member
246 of the vaginal microbiota, *L. crispatus,* isolated from a vaginally-born, 20-year-old

247 daughter and her mother (pair I) had highly similar genome sequences.

248 Other studies have compared the vaginal microbiotas of mothers and daughters
249 without detecting notable similarity between them [34, 35]. There are multiple reasons
250 that high similarity between mothers and daughters was not observed in these studies.
251 First, the effect of birth mode was not analyzed in these previous studies. In our study,
252 there was a trend toward greater community similarity within mother/daughter pairs,
253 regardless of birth mode, than between unrelated subjects in different mother/daughter
254 pairs (Figure 2A). However, most (10/13) of the daughters in our study were born by
255 vaginal delivery. Indeed, when mother/daughter pairs were separated by the daughter's
256 birth mode, the vaginal microbiotas of mothers and their daughters were significantly
257 different if the daughter was born by C- section. So, if many of the daughters in the
258 other studies were born by C-section then high similarity between mothers and
259 daughters would not be expected. With C- section rates of ~30% in the United States
260 (study site for [35]) and ~36% in South Korea (study site for [34]) this is a possibility [36,
261 37]. Additionally, our study focused on adolescent daughters (age 15-21) while the other
262 studies focused on either younger or older daughters. Since reproductive stage seems
263 to influence the structure of the vaginal microbiota [38], differences in reproductive
264 stage may contribute to differences in vaginal community composition between mothers
265 and daughters. Finally, we used a different method of comparing the vaginal
266 microbiotas of mothers and daughters. We calculated distances between mothers and
267 daughter using θ_{YC}, a metric that accounts for the relative abundances of shared and
268 non-shared OTUs, while the other studies were based on community types [35] and
269 Unifrac [34]. Although an overall community similarity was not observed in these

270 studies, specific community members (*Lactobacillus* and *Prevotella*) were identified as
271 most heritable in one study [34].

272 Based on the number of SNPs observed between the mother and daughter *L.*
273 *crispatus* isolates and published mutations rates for *L. casei* Zhang [33], we estimated
274 that *L. crispatus* would have an *in vivo* doubling time of 13.2-20.6 hours, depending on
275 the specific isolates compared. The doubling time estimates of 13.2 hours and
276 20.6 hours for *L. crispatus in vivo* are within the range estimated for other bacteria in
277 their natural environments, including *Escherichia coli* (15 hours) and *Salmonella*
278 *enterica* (25 hours) [39]. These doubling times are faster than the 4.1-5.6 days doubling
279 times measured for *L. casei* Shirota in mouse intestines, where its growth rate was
280 insufficient to maintain colonization [40]. Although the actual growth and mutation rates
281 of *L. crispatus* in the human vagina have not been measured, we estimated reasonable
282 *in vivo* doubling times for vaginal *L. crispatus* based on the observed number of SNPs
283 between *L. crispatus* isolates from mother/daughter pair I, the age of daughter I and *L.*
284 *casei* Zhang mutation rates. Considering the uncertainty in the estimates, transmission
285 of *L. crispatus* from mother to daughter at birth followed by the accumulation of
286 independent mutations during 20 years of persistence in the mother and daughter is a
287 plausible explanation for the observed recombination-filtered SNPs. Future studies
288 comparing genomes of *L. crispatus* isolates from more mother/daughter pairs with a
289 variety of daughter ages are needed.

290 The 2 *L. crispatus* isolates from the mother had highly similar genomes, differing
291 by only 11 recombination-filtered SNPs. A previous study also observed high similarity
292 between the genomes of multiple vaginal *L. crispatus* isolates from one individual, noting

293 that they were indistinguishable [41]. Future investigations of *L. crispatus* genomic
294 variation within an individual may yield further insight on colonization and dynamics of the
295 vaginal microbiota.

296 Consistent with a previous study, *L. crispatus* isolates from the human vagina were
297 phylogenetically intermixed with isolates from the human urinary tract, including highly
298 similar vaginal (ERS1867668 (SAMEA104208650)) and bladder (ERS1867667
299 (SAMEA104208649)) isolates from the same subject (Figure 3) [42].

300 The health implications of vertical transmission of the vaginal microbiota are
301 unknown and were not addressed in this study. However, because vertical transmission
302 seems to be an important factor in determining the composition of the vaginal
303 microbiota there may be important consequences. Vertical transmission of the vaginal
304 microbiota may be one mechanism for maintaining human microbiota over generations
305 via a consistent and specific seeding of the newborn microbiota. Delivery mode is an
306 important factor in determining the early composition of the gut microbiota [43, 44] and
307 is a risk factor for development of immune-related disorders later in life [45]. This
308 suggests an important role for the mother's vaginal microbiota in seeding the infant and
309 setting the stage for development of the gut microbiota. Therefore, maintenance of the
310 vaginal microbiota between generations may be critical for gut microbiota development
311 in each generation.

312 Additionally, the vaginal microbiota plays an important if not well understood role
313 in reproductive health, with associations between vaginal microbiota composition and
314 infection susceptibility, BV and preterm birth [46-48]. Evidence from this study suggests
315 that transmission of microbes from mother to daughter at birth may influence the

316 composition of the daughter's microbiota later in life and may contribute to the
317 maintenance of specific members of the human vaginal microbiota over generations.
318 This study provides tantalizing evidence of vertical transmission of the vaginal
319 microbiota. However, this was a small study with only 13 mother/daughter pairs (92%
320 white) and 3/13 daughters born by C-section. Genomic analysis of isolates was limited
321 to one member of the vaginal microbiota from 1 mother/daughter pair. Future studies in
322 larger populations, including more racially diverse subjects, more daughters born by C-
323 section and analysis of more isolate genome sequences or metagenomes are required
324 to validate these findings.

325 **Conclusions**

326 At least some members of the vaginal microbiota are transmitted from mother to
327 daughter at birth and persist into adolescence. Birth-mode dependent vaginal
328 microbiota transmission is an important factor determining vaginal microbiota
329 composition.

330 **Figures**

331 **Figure 1. Vaginal bacterial community compositions of mother/daughter pairs.**
332 Relative abundances of OTUs in weekly vaginal swab samples from 13
333 mother/daughter pairs. Mother/daughter pairs were ordered by average within pair θ_{YC}
334 distances, with the most similar pair (I) on top and the least similar pair (XIII) on the
335 bottom. OTUs with a minimum of 200 sequences in the dataset overall and present at a
336 relative abundance greater than 2% in at least 1 sample were included in the heat map.

337 **Figure 2. Average distances between vaginal bacterial communities.** A. Average
338 θ_{YC} distances between subjects from different mother/daughter pairs (between pairs),

339 between subjects within a mother/daughter pair (within pair) and between samples from
340 the same subject (within subject). P-values for comparisons that were significantly
341 different by Dunn's posttest are shown (Kruskal-Wallis p-value= 8.154e-10). B. Average
342 θ_{YC} distances between subjects within a mother/daughter pair for daughters born by
343 vaginal birth and by C-section. Wilcoxon (Mann-Whitney) test p-value is shown. In the
344 box and whiskers plots, the median θ_{YC} distance is indicated by a line, values within the
345 first to the third quartiles are inside the box and the whiskers extend to the smallest and
346 largest values within 1.5x the interquartile range.

347 **Figure 3. Phylogenetic relationships between *L. crispatus* strains.** Maximum
348 likelihood tree based on recombination-filtered SNP distances between *L. crispatus*
349 genome sequences of isolates from mother/daughter pair I and other *L. crispatus*
350 strains with publicly available genomes. Tip labels indicate *L. crispatus* strain names
351 and NCBI BioSample identifiers. Bootstrap values were greater than or equal to 0.65.

352 **Supplemental Figure 1. Principal coordinates analysis (PCoA) of vaginal**
353 **microbiota from 13 mother/daughter pairs.** The θ_{YC} distances between 101 vaginal
354 microbiota samples are represented by PCoA. Samples from daughters are represented
355 by triangles and samples from mothers by circles. Each mother/daughter pair is
356 represented by a unique color. Biplot arrows represent the 3 OTUs most correlated with
357 position on the PCoA plot.

358 **List of abbreviations**

359 C-section: Cesarean section

360 rRNA: ribosomal RNA

361 OTU: operational taxonomic unit

362 SNPs: single nucleotide polymorphisms

363 PCoA: principal coordinates analysis

364 **Declarations**

365 - Ethics approval and consent to participate

366 All subjects provided written informed consent. The study was approved by the

367 University of Michigan IRB (HUM00086661).

368 - Consent for publication

369 Not applicable.

370 - Availability of data and material

371 The raw sequence data generated in this study are available in the NCBI's SRA:

372 Bacterial 16S rRNA gene sequences: BioProject PRJNA547595

373 *L. crispatus* draft genome sequences: BioProject PRJNA547620

374 GitHub repository:

375 https://github.com/cbassis/MotherDaughter_Vaginal_Microbiota.study

376 This repository includes:

377 - the mothur batch file with steps used to process and analyze 16S rRNA gene

378 sequences

379 - mothur output files used in final bacterial community analysis and figures

380 - R code for manuscript figures, statistics and genomic analysis

381 - Competing interests

382 The authors declare that they have no competing interests.

383 - Funding

384 Not applicable.

385 • Authors' contributions

386 CMB was involved in study design and planning, data analysis, figure preparation
387 and manuscript writing. KAB was involved in subject recruitment, sample
388 processing, isolation of *L. crispatus* genomic DNA for sequencing, data analysis and
389 manuscript editing. DES was involved in subject recruitment, sample processing and
390 manuscript editing. KS was involved in genomic data analysis, interpretation of
391 genomic data, phylogenetic tree construction and manuscript editing. AP was
392 involved in genomic data analysis and manuscript editing. ES was involved in
393 genomic data analysis and interpretation and manuscript review. VIA was involved in
394 subject recruitment and planning. EHQ was involved in study design and planning,
395 subject recruitment and manuscript editing. VBY was involved in study design and
396 manuscript editing. JDB was involved with study design and planning, subject
397 recruitment and manuscript editing. All authors read and approved the final
398 manuscript.

399 • Acknowledgements

400 We'd like to thank the participants who contributed samples. We'd also like to thank
401 Megan Lemieur for help with study organization and Qualtrics survey implementation
402 and the University of Michigan Microbial Systems Molecular Biology Laboratory for
403 swab DNA isolation, bacterial community and genomic library preparation and
404 sequencing. We'd also like to thank Ingrid Bergin and members of the Young lab for
405 many discussions related to this project over the years. This research was also
406 supported in part through computational resources and services provided by
407 Advanced Research Computing at the University of Michigan, Ann Arbor.

408 ## References

409 1. Callahan BJ, DiGiulio DB, Goltsman DSA, Sun CL, Costello EK, Jeganathan P, Biggio JR,
410 Wong RJ, Druzin ML, Shaw GM, et al: **Replication and refinement of a vaginal microbial**
411 **signature of preterm birth in two racially distinct cohorts of US women.** *Proceedings of*
412 *the National Academy of Sciences* 2017.
413 2. Klatt NR, Cheu R, Birse K, Zevin AS, Perner M, Noël-Romas L, Grobler A, Westmacott G,
414 Xie IY, Butler J, et al: **Vaginal bacteria modify HIV tenofovir microbicide efficacy in**
415 **African women.** *Science (Wash DC)* 2017, **356:**938.
416 3. McClelland RS, Lingappa JR, Srinivasan S, Kinuthia J, John-Stewart GC, Jaoko W,
417 Richardson BA, Yuhas K, Fiedler TL, Mandaliya KN, et al: **Evaluation of the association**
418 **between the concentrations of key vaginal bacteria and the increased risk of HIV**
419 **acquisition in African women from five cohorts: a nested case-control study.** *The*
420 *Lancet Infectious Diseases* 2018, **18:**554-564.
421 4. Durack J, Lynch SV: **The gut microbiome: Relationships with disease and opportunities**
422 **for therapy.** *The Journal of Experimental Medicine* 2018**:**jem.20180448.
423 5. **Structure, function and diversity of the healthy human microbiome.** *Nature* 2012,
424 **486:**207-214.
425 6. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, Karlebach S, Gorle R,
426 Russell J, Tacket CO, et al: **Vaginal microbiome of reproductive-age women.**
427 *Proceedings of the National Academy of Sciences* 2011, **108:**4680-4687.
428 7. Bassis CM, Allsworth JE, Wahl HN, Sack DE, Young VB, Bell JD: **Effects of intrauterine**
429 **contraception on the vaginal microbiota.** *Contraception* 2017, **96:**189-195.
430 8. O'Hanlon DE, Moench TR, Cone RA: **Vaginal pH and Microbicidal Lactic Acid When**
431 **Lactobacilli Dominate the Microbiota.** *PLoS ONE* 2013, **8:**e80074.
432 9. Tachedjian G, O'Hanlon DE, Ravel J: **The implausible "in vivo" role of hydrogen**
433 **peroxide as an antimicrobial factor produced by vaginal microbiota.** *Microbiome* 2018,
434 **6:**29.
435 10. Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, Knight R:
436 **Delivery mode shapes the acquisition and structure of the initial microbiota across**
437 **multiple body habitats in newborns.** *Proceedings of the National Academy of Sciences*
438 2010, **107:**11971-11975.
439 11. Madan JC, Hoen AG, Lundgren SN, et al.: **Association of cesarean delivery and formula**
440 **supplementation with the intestinal microbiome of 6-week-old infants.** *JAMA*
441 *Pediatrics* 2016, **170:**212-219.
442 12. Seekatz AM, Theriot CM, Molloy CT, Wozniak KL, Bergin IL, Young VB: **Fecal Microbiota**
443 **Transplantation Eliminates *Clostridium difficile* in a Murine Model of Relapsing**
444 **Disease.** *Infection and Immunity* 2015, **83:**3838-3846.
445 13. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD: **Development of a Dual-**
446 **Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence**
447 **Data on the MiSeq Illumina Sequencing Platform.** *Appl Environ Microbiol* 2013,
448 **79:**5112-5120.
449 14. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA,
450 Oakley BB, Parks DH, Robinson CJ, et al: **Introducing mothur: Open-Source, Platform-**

451 **Independent, Community-Supported Software for Describing and Comparing**
452 **Microbial Communities.** *Appl Environ Microbiol* 2009, **75:**7537-7541.
453 15. Schloss PD: **A High-Throughput DNA Sequence Aligner for Microbial Ecology Studies.**
454 *PLoS ONE* 2009, **4:**e8230.
455 16. Westcott SL, Schloss PD: **De novo clustering methods outperform reference-based**
456 **methods for assigning 16S rRNA gene sequences to operational taxonomic units.** *PeerJ*
457 2015, **3:**e1487.
458 17. Schloss PD, Westcott SL: **Assessing and Improving Methods Used in Operational**
459 **Taxonomic Unit-Based Approaches for 16S rRNA Gene Sequence Analysis.** *Appl Environ*
460 *Microbiol* 2011, **77:**3219-3226.
461 18. Wang Q, Garrity GM, Tiedje JM, Cole JR: **Naive Bayesian Classifier for Rapid Assignment**
462 **of rRNA Sequences into the New Bacterial Taxonomy.** *Appl Environ Microbiol* 2007,
463 **73:**5261-5267.
464 19. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske
465 CR, Tiedje JM: **Ribosomal Database Project: data and tools for high throughput rRNA**
466 **analysis.** *Nucleic Acids Res* 2014, **42:**D633-642.
467 20. Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schäffer AA: **Database**
468 **indexing for production MegaBLAST searches.** *Bioinformatics (Oxf)* 2008, **24:**1757-1764.
469 21. Yue JC, Clayton MK: **A similarity measure based on species proportions.**
470 *Communications in Statistics-Theory and Methods* 2005, **34:**2123-2131.
471 22. **FastQC** [http://www.bioinformatics.babraham.ac.uk/projects/fastqc/]
472 23. Bolger AM, Lohse M, Usadel B: **Trimmomatic: a flexible trimmer for Illumina sequence**
473 **data.** *Bioinformatics (Oxf)* 2014, **30:**2114-2120.
474 24. Li H, Durbin R: **Fast and accurate short read alignment with Burrows-Wheeler**
475 **transform.** *Bioinformatics (Oxf)* 2009, **25:**1754-1760.
476 25. **BWA** [http://bio-bwa.sourceforge.net]
477 26. **Picard** [https://broadinstitute.github.io/picard/]
478 27. **Calling SNPs/INDELs with SAMtools/BCFtools**
479 [http://samtools.sourceforge.net/mpileup.shtml]
480 28. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K,
481 Altshuler D, Gabriel S, Daly M, DePristo MA: **The Genome Analysis Toolkit: a**
482 **MapReduce framework for analyzing next-generation DNA sequencing data.** *Genome*
483 *Res* 2010, **20:**1297-1303.
484 29. Pon A, Marcu A, Arndt D, Grant JR, Sajed T, Liang Y, Wishart DS: **PHASTER: a better,**
485 **faster version of the PHAST phage search tool.** *Nucleic Acids Res* 2016, **44:**W16-W21.
486 30. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL:
487 **Versatile and open software for comparing large genomes.** *Genome Biol* 2004, **5:**R12.
488 31. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, Parkhill J, Harris SR:
489 **Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome**
490 **sequences using Gubbins.** *Nucleic Acids Res* 2015, **43:**e15.
491 32. Stamatakis A: **RAxML version 8: a tool for phylogenetic analysis and post-analysis of**
492 **large phylogenies.** *Bioinformatics (Oxford, England)* 2014, **30:**1312-1313.

493 33. Wang J, Dong X, Shao Y, Guo H, Pan L, Hui W, Kwok L-Y, Zhang H, Zhang W: **Genome**
494 **adaptive evolution of Lactobacillus casei under long-term antibiotic selection**
495 **pressures.** *BMC Genomics* 2017, **18:**320.
496 34. Si J, You HJ, Yu J, Sung J, Ko G: ***Prevotella* as a Hub for Vaginal Microbiota under the**
497 **Influence of Host Genetics and Their Association with Obesity.** *Cell Host & Microbe*
498 2016, **21:**97-105.
499 35. Hickey RJ, Zhou X, Settles ML, Erb J, Malone K, Hansmann MA, Shew ML, Van Der Pol B,
500 Fortenberry JD, Forney LJ: **Vaginal Microbiota of Adolescent Girls Prior to the Onset of**
501 **Menarche Resemble Those of Reproductive-Age Women.** *mBio* 2015, **6:**e00097-00015.
502 36. Kim SJ, Kim SJ, Han K-T, Park E-C: **Medical costs, Cesarean delivery rates, and length of**
503 **stay in specialty hospitals vs. non-specialty hospitals in South Korea.** *PLOS ONE* 2017,
504 **12:**e0188612.
505 37. Caughey AB, Cahill AG, Guise J-M, Rouse DJ: **Safe prevention of the primary cesarean**
506 **delivery.** *American Journal of Obstetrics and Gynecology* 2014, **210:**179-193.
507 38. Brotman RM, Shardell MD, Gajer P, Fadrosh D, Chang K, Silver MI, Viscidi RP, Burke AE,
508 Ravel J, Gravitt PE: **Association between the vaginal microbiota, menopause status,**
509 **and signs of vulvovaginal atrophy.** *Menopause* 2018, **25:**1321-1330.
510 39. Gibson B, Wilson Daniel J, Feil E, Eyre-Walker A: **The distribution of bacterial doubling**
511 **times in the wild.** *Proceedings of the Royal Society B: Biological Sciences* 2018,
512 **285:**20180789.
513 40. Lee YK, Ho PS, Low CS, Arvilommi H, Salminen S: **Permanent Colonization by**
514 ***Lactobacillus casei* Is Hindered by the Low Rate of Cell Division in Mouse Gut.** *Appl*
515 *Environ Microbiol* 2004, **70:**670.
516 41. Abdelmaksoud AA, Koparde VN, Sheth NU, Serrano MG, Glascock AL, Fettweis JM,
517 Strauss JF, 3rd, Buck GA, Jefferson KK: **Comparison of Lactobacillus crispatus isolates**
518 **from Lactobacillus-dominated vaginal microbiomes with isolates from microbiomes**
519 **containing bacterial vaginosis-associated bacteria.** *Microbiology* 2016, **162:**466-475.
520 42. Thomas-White K, Forster SC, Kumar N, Van Kuiken M, Putonti C, Stares MD, Hilt EE, Price
521 TK, Wolfe AJ, Lawley TD: **Culturing of female bladder bacteria reveals an**
522 **interconnected urogenital microbiota.** *Nat Commun* 2018, **9:**1557.
523 43. Wampach L, Heintz-Buschart A, Fritz JV, Ramiro-Garcia J, Habier J, Herold M,
524 Narayanasamy S, Kaysen A, Hogan AH, Bindl L, et al: **Birth mode is associated with**
525 **earliest strain-conferred gut microbiome functions and immunostimulatory potential.**
526 *Nature Communications* 2018, **9:**5091.
527 44. Rutayisire E, Huang K, Liu Y, Tao F: **The mode of delivery affects the diversity and**
528 **colonization pattern of the gut microbiota during the first year of infants' life: a**
529 **systematic review.** *BMC Gastroenterology* 2016, **16:**86.
530 45. Tamburini S, Shen N, Wu HC, Clemente JC: **The microbiome in early life: implications**
531 **for health outcomes.** *Nature Medicine* 2016, **22:**713.
532 46. Bayigga L, Kateete DP, Anderson DJ, Sekikubo M, Nakanjako D: **Diversity of vaginal**
533 **microbiota in sub-Saharan Africa and its effects on HIV transmission and prevention.**
534 *American Journal of Obstetrics and Gynecology* 2018.
535 47. Tamarelle J, Thiébaut ACM, de Barbeyrac B, Bébéar C, Ravel J, Delarocque-Astagneau E:
536 **The vaginal microbiota and its association with human papillomavirus, Chlamydia**

537 **trachomatis, Neisseria gonorrhoeae and Mycoplasma genitalium infections: a**
538 **systematic review and meta-analysis.** *Clinical Microbiology and Infection* 2019, **25:**35-
539 47.
540 48. Callahan BJ, DiGiulio DB, Goltsman DSA, Sun CL, Costello EK, Jeganathan P, Biggio JR,
541 Wong RJ, Druzin ML, Shaw GM, et al: **Replication and refinement of a vaginal microbial**
542 **signature of preterm birth in two racially distinct cohorts of US women.** *Proceedings of*
543 *the National Academy of Sciences* 2017, **114:**9966.
544

545

Figure 1

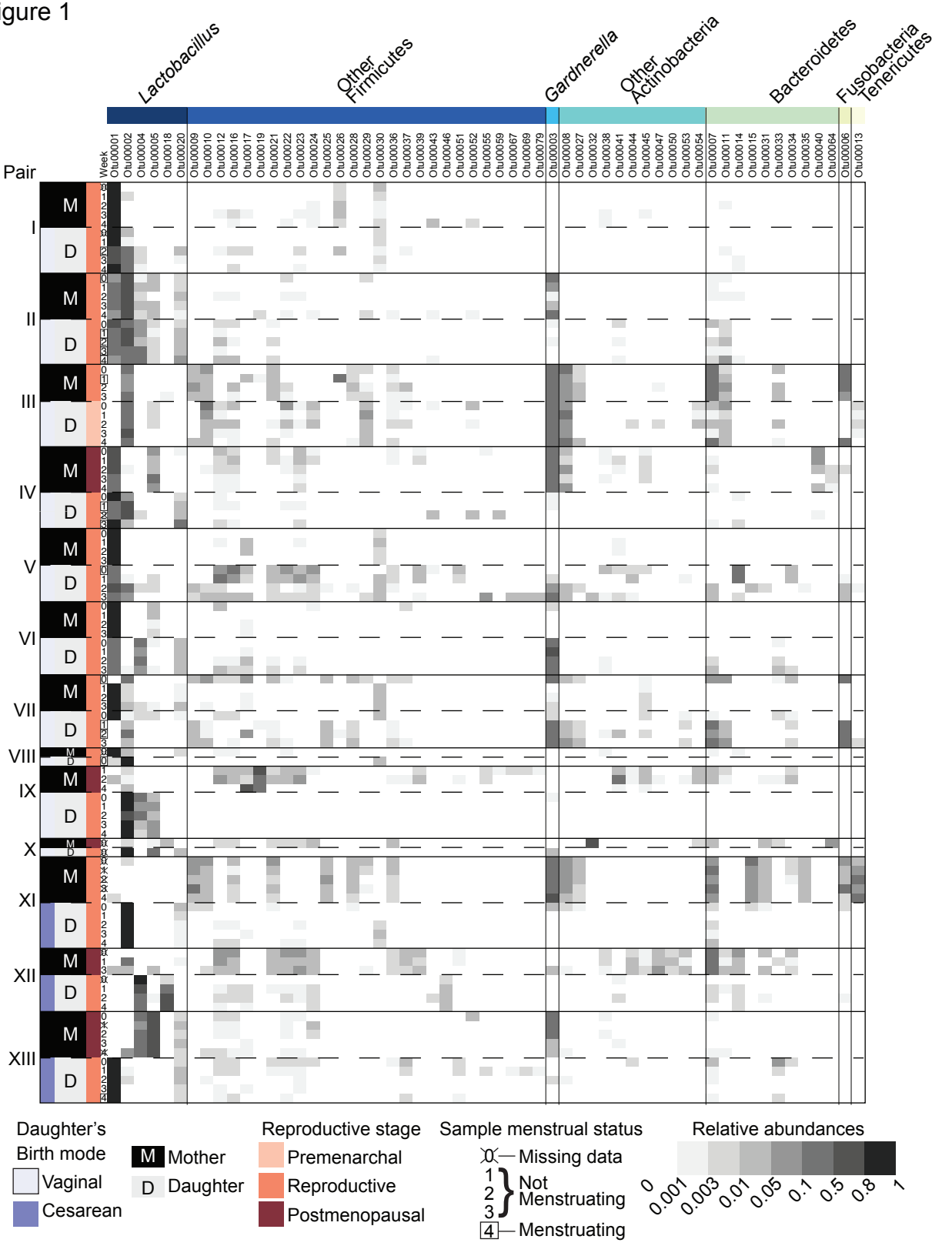

Lactobacillus
Other Firmicutes
Gardnerella
Other Actinobacteria
Bacteroidetes
Fusobacteria Tenericutes
Pair
Week
I
II
III
IV
V
VI
VII
VIII
IX
X
XI
XII
XIII
M
D
Daughter's Birth mode
Vaginal
Cesarean
M Mother
D Daughter
Reproductive stage
Premenarchal
Reproductive
Postmenopausal
Sample menstrual status
Missing data
1 2 3 Not Menstruating
4 Menstruating
Relative abundances
0
0.001
0.003
0.01
0.05
0.1
0.5
0.8
1

Figure 2

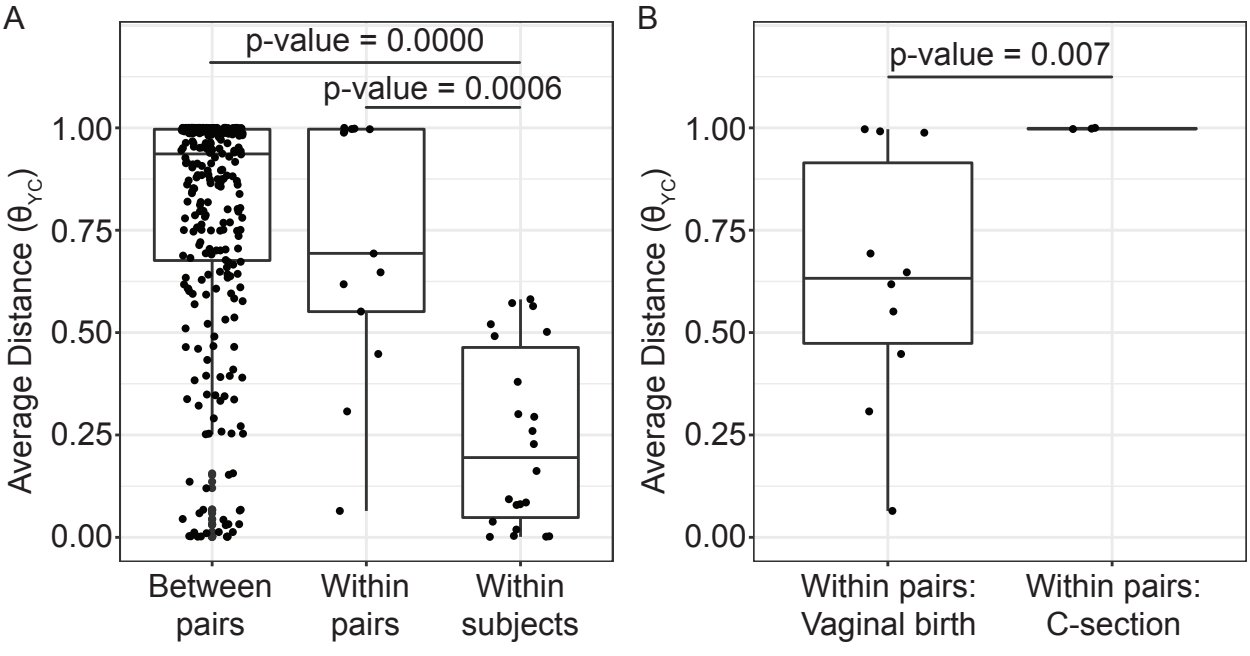
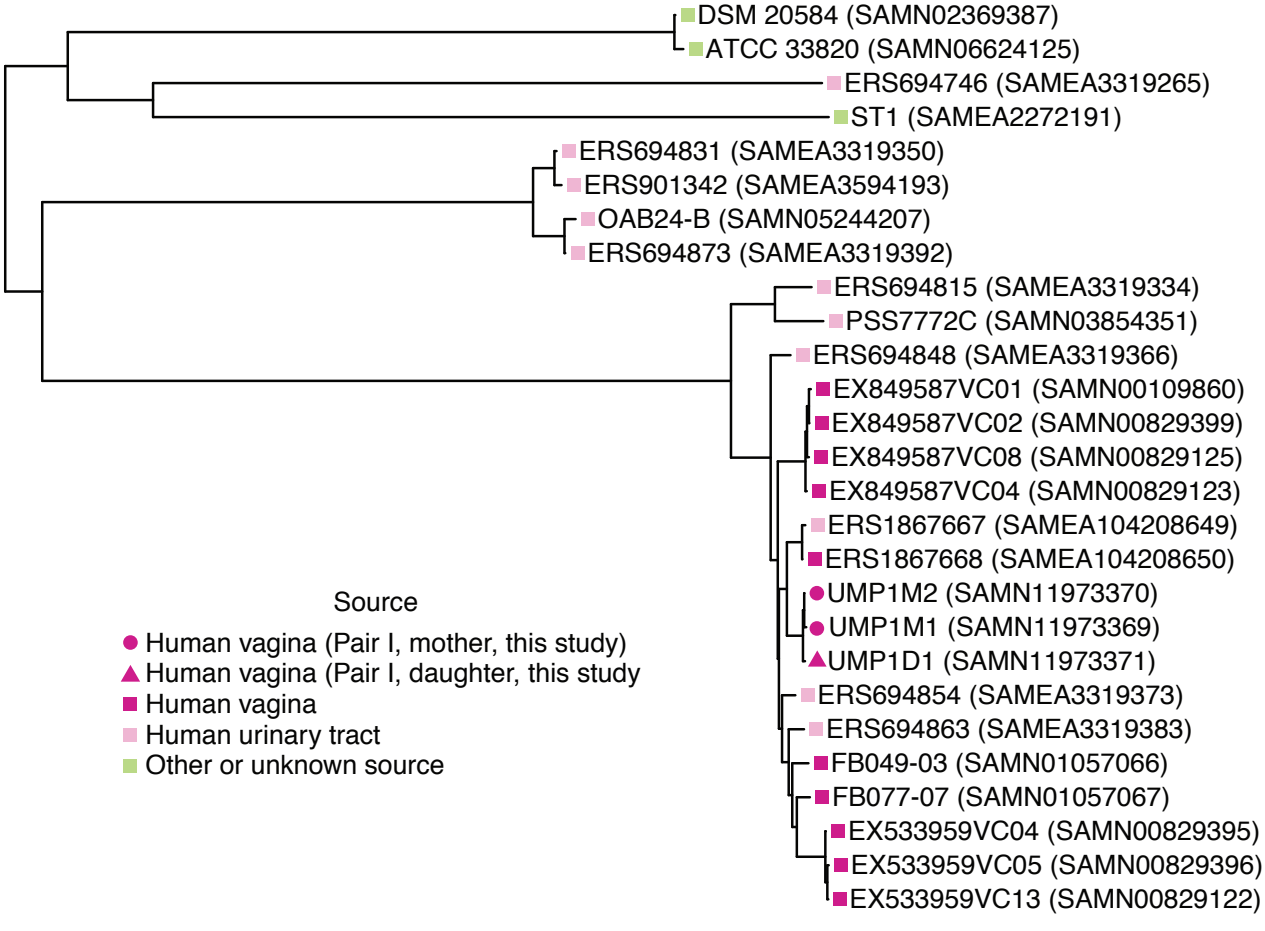
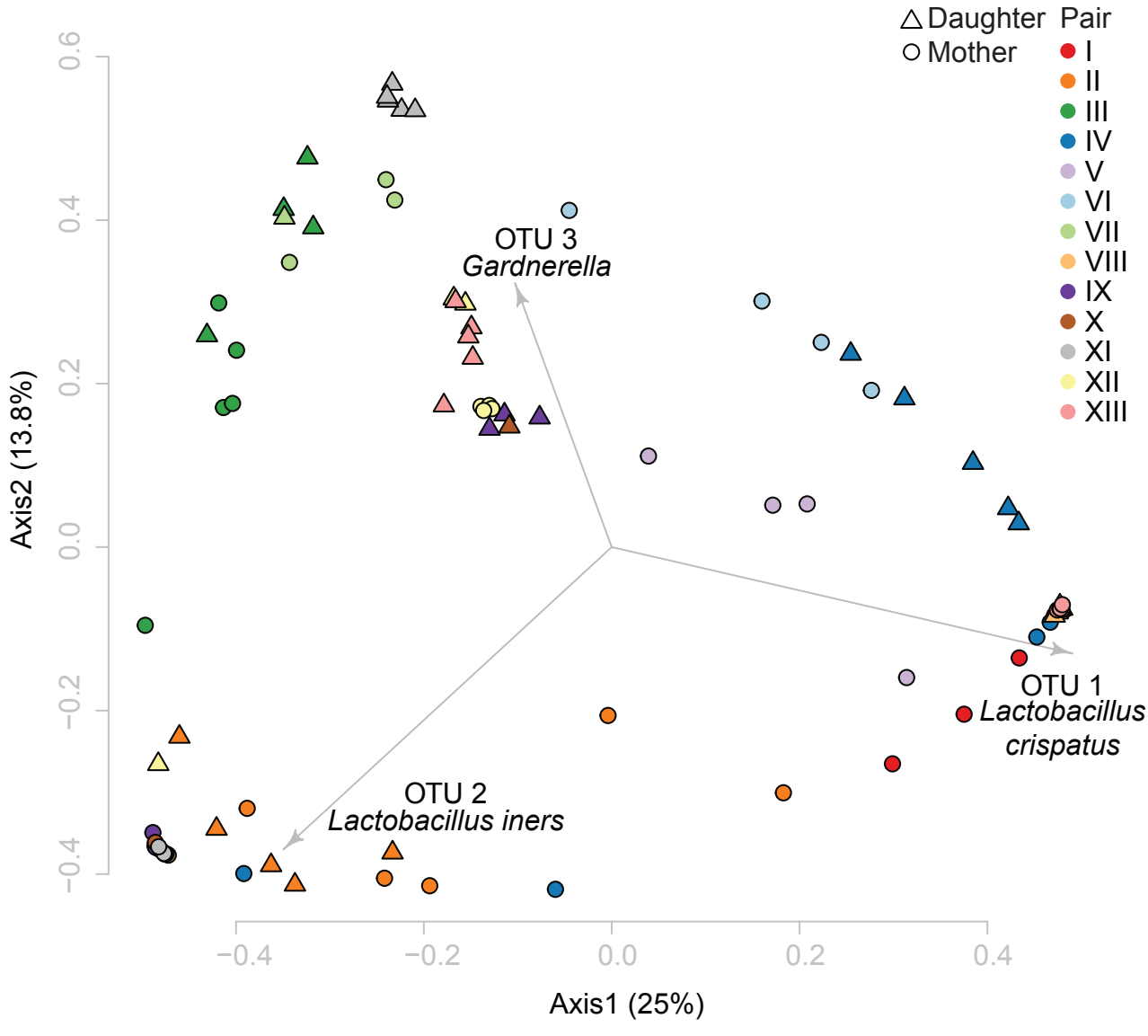

A
p-value = 0.0000
p-value = 0.0006
Average Distance (θ_{YC})
1.00
0.75
0.50
0.25
0.00
Between pairs
Within pairs
Within subjects
B
p-value = 0.007
Average Distance (θ_{YC})
1.00
0.75
0.50
0.25
0.00
Within pairs: Vaginal birth
Within pairs: C-section

Figure 3


DSM 20584 (SAMN02369387)
ATCC 33820 (SAMN06624125)
ERS694746 (SAMEA3319265)
ST1 (SAMEA2272191)
ERS694831 (SAMEA3319350)
ERS901342 (SAMEA3594193)
OAB24-B (SAMN05244207)
ERS694873 (SAMEA3319392)
ERS694815 (SAMEA3319334)
PSS7772C (SAMN03854351)
ERS694848 (SAMEA3319366)
EX849587VC01 (SAMN00109860)
EX849587VC02 (SAMN00829399)
EX849587VC08 (SAMN00829125)
EX849587VC04 (SAMN00829123)
ERS1867667 (SAMEA104208649)
ERS1867668 (SAMEA104208650)
UMP1M2 (SAMN11973370)
UMP1M1 (SAMN11973369)
UMP1D1 (SAMN11973371)
ERS694854 (SAMEA3319373)
ERS694863 (SAMEA3319383)
FB049-03 (SAMN01057066)
FB077-07 (SAMN01057067)
EX533959VC04 (SAMN00829395)
EX533959VC05 (SAMN00829396)
EX533959VC13 (SAMN00829122)
Source
Human vagina (Pair I, mother, this study)
Human vagina (Pair I, daughter, this study
Human vagina
Human urinary tract
Other or unknown source

Supplemental Figure 1