

Vaginal Microbiota of Adolescent Girls Prior to the Onset of Menarche Resemble Those of Reproductive-Age Women

Roxana J. Hickey,^{a,b} Xia Zhou,^{a*} Matthew L. Settles,^{a,b} Julie Erb,^c Kristin Malone,^c Melanie A. Hansmann,^c Marcia L. Shew,^d Barbara Van Der Pol,^e J. Dennis Fortenberry,^d Larry J. Forney^{a,b}

Department of Biological Sciences, University of Idaho, Moscow, Idaho, USA^a; Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, Idaho, USA^b; FemCare Division, Procter & Gamble, Cincinnati, Ohio, USA^c; Department of Pediatrics, Indiana University, Indianapolis, Indiana, USA^d; Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama, USA^e

* Present address: Xia Zhou, Vista Encology, Inc., Olympia, Washington, USA.

ABSTRACT Puberty is an important developmental stage wherein hormonal shifts mediate the physical and physiological changes that lead to menarche, but until now, the bacterial composition of vaginal microbiota during this period has been poorly characterized. We performed a prospective longitudinal study of perimenarcheal girls to gain insight into the timing and sequence of changes that occur in the vaginal and vulvar microbiota during puberty. The study enrolled 31 healthy, premenarcheal girls between the ages of 10 and 12 years and collected vaginal and vulvar swabs quarterly for up to 3 years. Bacterial composition was characterized by Roche 454 pyrosequencing and classification of regions V1 to V3 of 16S rRNA genes. Contrary to expectations, lactic acid bacteria, primarily *Lactobacillus* spp., were dominant in the microbiota of most girls well before the onset of menarche in the early to middle stages of puberty. *Gardnerella vaginalis* was detected at appreciable levels in approximately one-third of subjects, a notable finding considering that this organism is commonly associated with bacterial vaginosis in adults. Vulvar microbiota closely resembled vaginal microbiota but often exhibited additional taxa typically associated with skin microbiota. Our findings suggest that the vaginal microbiota of girls begin to resemble those of adults well before the onset of menarche.

IMPORTANCE This study addresses longitudinal changes in vaginal and vulvar microbial communities prior to and immediately following menarche. The research is significant because microbial ecology of the vagina is an integral aspect of health, including resistance to infections. The physiologic changes of puberty and initiation of cyclic menstruation are likely to have profound effects on vaginal microbiota, but almost nothing is known about changes that normally occur during this time. Our understanding has been especially hampered by the lack of thorough characterization of microbial communities using techniques that do not rely on the cultivation of fastidious bacteria, as well as a dearth of studies on girls in the early to middle stages of puberty. This study improves our understanding of the normal development of vaginal microbiota during puberty and onset of menarche and may better inform clinical approaches to vulvovaginal care of adolescent girls.

Received 2 February 2015 Accepted 18 February 2015 Published 24 March 2015

Citation Hickey RJ, Zhou X, Settles ML, Erb J, Malone K, Hansmann MA, Shew ML, Van Der Pol B, Fortenberry JD, Forney LJ. 2015. Vaginal microbiota of adolescent girls prior to the onset of menarche resemble those of reproductive-age women. mBio 6(2):e00097-15. doi:10.1128/mBio.00097-15.

Editor Vanessa Sperandio, University of Texas Southwestern Medical Center

Copyright © 2015 Hickey et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution-Noncommercial-ShareAlike 3.0 Unported license](https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

Address correspondence to J. Dennis Fortenberry, jfortenb@iu.edu, or Larry J. Forney, lforney@uidaho.edu.

This article is a direct contribution from a Fellow of the American Academy of Microbiology.

Understanding changes in vaginal bacterial communities over a woman's life span is essential to comprehending normal development, physiological function and health, and susceptibility to disease. Until now, vaginal microbiota before puberty were thought to be relatively stable assemblages of aerobic, anaerobic, and enteric bacterial populations (1–4). After menarche, the vaginal microbiota of healthy adults are typified by high numbers of homofermentative lactic acid bacteria, which contribute to acidification of the vaginal microenvironment through the production of lactate and other organic acids (5, 6). Various species of *Lactobacillus* have been identified as the predominant lactic acid bacteria in most adult women, and the ecological function of lactate production is further conserved by genera such as *Streptococcus*

and *Atopobium*, which are found in a subset of women (7). Recent studies using cultivation-independent methods have revealed the substantial complexity and temporal variability of the vaginal microbiota (8–10). Multiple community types distinguished by differences in the kinds and relative abundances of bacterial populations present are consistently found among healthy adult women (7, 8, 11, 12). These findings imply the absence of a “core” healthy vaginal microbiota and underscore the importance of delineating differences in community composition among individuals as well as changes over time.

To date, most studies of the vaginal microbiota have focused exclusively on reproductive-age women. As a result, little is known about when and how these communities are established during

TABLE 1 Characteristics of adolescent study participants

Subject race and ethnicity	No. (%)		Who achieved menarche during study	Who had a participating mother	Mean age at enrollment (yr)	Mean duration of study participation (yr)
	Total					
All subjects	31 (100)		21 (67.7)	24 (77.4)	10.9	1.6
Black, non-Hispanic	21 (67.7)		14 (66.7)	17 (81.0)	10.9	1.7
White, non-Hispanic	7 (22.6)		5 (71.4)	4 (57.1)	10.8	1.4
White, Hispanic	1 (3.2)		0 (0.0)	1 (100)	10.2	1.4
Native American, non-Hispanic	2 (6.5)		2 (100)	2 (100)	11.5	1.3

puberty. Changes in the composition and function of the vaginal microbiota during puberty are thought to be mediated by estrogen-stimulated glycogen production in the vaginal epithelium (6), but the timing and sequence of events during this period are not well studied. Menarche itself does not signal the completion of puberty, and pubertal hormonal influences on vaginal microbial communities may continue for months or even years following menarche. Using cultivation-dependent methods or microscopic examination, past studies of premenarcheal vaginal microbiota found bacterial communities with low numbers of strict and facultative anaerobes, with most species of apparently enteric origin (2–4). *Lactobacillus* species were rarely observed and, when found, constituted only a minor proportion of the total bacteria. Transition to adult-like vaginal microbial communities is not well documented but apparently occurs over a short time, as the vaginal microbiota of perimenarcheal and postmenarcheal 13- to 18-year-olds were found to resemble those of older women (13–15). However, most past studies are limited by inherent biases imposed by cultivation-dependent methods, which fail to account for many bacterial taxa. Furthermore, we are unaware of studies that specifically characterized community composition in detail while evaluating subsequent physical and physiological changes through menarche and thereafter. This lack of data highlights the need for longitudinal characterization of the vaginal microbial communities in perimenarcheal girls (i.e., before, during, and following menarche).

There are several reasons to pursue a better understanding of the perimenarcheal vaginal microbiota. Clinically, vulvar and vaginal complaints such as vulvovaginitis are common among premenarcheal girls and are often ascribed to poor hygiene or physiologic leukorrhea (vaginal discharge due to estrogen stimulation) (16–18). Numerous studies have reported bacterial vaginosis in adolescent girls, using diagnostic criteria developed for adult women (13, 19–24). Without a frame of reference for “normal” vaginal microbiota in healthy adolescents, the clinical relevance of microbiota resembling that associated with bacterial vaginosis is uncertain. Furthermore, as girls progress into menarche, menstrual hygiene behaviors, including use of menstrual pads and tampons, bathing habits, and douching, may alter existing vaginal microbiota (20, 25–30). Finally, changes in the early vaginal microbiota may have lasting influences on subsequent vaginal health, but our understanding of the complex interactions of immune tolerance of indigenous bacterial populations, immune surveillance for vaginal pathogens, variability in vaginal microbiota, and reproductive health outcomes remains primitive (31, 32).

To better understand changes in both the vaginal and vulvar microbiota before, during, and after menarche, 31 healthy pre-

menarcheal girls were enrolled between 10 and 12 years of age in a prospective longitudinal study in which girls were sampled quarterly for up to 3 years. The bacterial community composition of the vaginal and vulvar microbiota of girls, and vaginal microbiota of a subsample of their mothers, in this prospectively followed cohort are described. Our findings suggest that the vaginal microbiota of adolescents resemble those of reproductive-age women, although not necessarily those of their mothers, well before the onset of menarche in early stages of pubertal development. Familiar bacterial species associated with the vaginal microbiota of adults were commonly found in girls, including *Lactobacillus crispatus*, *L. iners*, *L. gasseri*, *L. jensenii*, and, notably, *Gardnerella vaginalis*. Following menarche, vaginal pH often remained above what is considered typical in healthy adult women even when lactobacilli were present in high proportions, raising the possibility that total bacterial loads may not reach levels seen in adults until later in puberty. These analyses provide the first detailed investigation of the progression of changes that occur in the vaginal microbiota over time during puberty.

RESULTS

Clinical study and data collection. A prospective longitudinal study was conducted in Indianapolis, IN, from June 2009 through June 2012 to assess changes in the vaginal and vulvar microbiota of perimenarcheal girls as they transitioned through menarche. Thirty-one healthy, asymptomatic girls representing black (*n* = 21), white/non-Hispanic (*n* = 7), white/Hispanic (*n* = 1) and Native American (*n* = 2) racial/ethnic groups were enrolled between June 2009 and August 2011. None of the girls had a history of sexual contact or recent antibiotic use. At enrollment, girls were between the ages of 10.0 and 12.9 years (mean, 10.9 years) and premenarcheal. The age range was selected because it represents a developmentally meaningful interval during which pubertal development typically begins (33, 34), and we aimed to increase the chances that girls would reach menarche and experience subsequent menstrual cycling while participating in the study. Twenty-one girls (67.7%) reached menarche during the study. In addition, the mothers of 24 girls participated by providing vaginal swabs annually. Biological maternity was not recorded or required for participation. Although age, race, and ethnicity of the mothers were captured, no analyses were performed on these data. Characteristics of the study participants are summarized in Table 1, with additional details included in Table S1 in the supplemental material. Both parental and adolescent consent were obtained at the time of enrollment.

Girls returned for sample collection and clinical examination at 3-month intervals for up to the duration of the 3-year study

(mean participation, 1.6 years; range, 1 day to 3.0 years). At each visit, up to three vaginal swabs and two vulvar swabs were collected by a female clinician as permitted by each girl. Breast and pubic development were assessed at each visit using Tanner's criteria (33), and vaginal pH was determined using commercial pH paper. Mothers provided self-collected vaginal swabs on an annual basis. A total of 457 swabs were processed for Roche 454 pyrosequencing of the hypervariable regions V1 to V3 of 16S rRNA genes: 198 vaginal swabs and 212 vulvar swabs from girls and 47 vaginal swabs from mothers. A summary of samples collected at each visit (with Tanner stage and menarche status indicated) is shown in Fig. S1 in the supplemental material. In total, there were 186 pairs of matched vagina-vulva samples from girls. Associated vaginal pH measurements were available for 65.7% of the sampling times from girls, while Nugent scores were obtained at only 24.8% of visits. We therefore elected not to analyze the Nugent score data.

16S rRNA sequencing and taxonomic assignment of reads.

Bacterial composition was determined based on sequencing the hypervariable regions V1 to V3 of 16S rRNA genes by Roche 454 pyrosequencing. High-quality reads were obtained from all but one vulvar swab sample, which was subsequently excluded from further analysis. The remaining 456 samples had a minimum of 462 and maximum of 31,569 reads after preprocessing (mean, 4,723 reads; median, 3,876 reads). Following taxonomic assignment of reads and summarization of the taxon abundance table as described in Materials and Methods, 78 taxa were identified to the species ($n = 9$), genus ($n = 60$), family ($n = 7$), or order ($n = 2$) level.

Clustering analyses of the perimenarcheal vaginal microbiota. The primary objectives of this study were to characterize the composition of vaginal microbiota in perimenarcheal girls, identify major community types, and assess similarities and differences in relation to menarche and pubertal development. Analyses were performed in R (35) using custom scripts available on GitHub.

To define the vaginal community types present in both girls and mothers in the study, hierarchical clustering was performed on the Bray-Curtis dissimilarity matrix calculated from Hellinger-standardized taxon abundance data. The average linkage method was identified as the best clustering method for the data using the minimum Gower distance (36), and seven clusters were selected as optimal using the silhouette method (37). Figure 1 shows the hierarchical clustering and community composition of all 198 vaginal swabs from girls and 47 vaginal swabs from mothers. Four clusters were characterized by high proportions of different *Lactobacillus* spp., including *L. crispatus* (cluster LC; $n = 87$), *L. iners* (LI; $n = 71$), *L. gasseri* (LG; $n = 22$), and *L. jensenii* (LJ; $n = 8$). A fifth cluster was characterized by high proportions of *Gardnerella vaginalis* (cluster GV; $n = 30$). The sixth cluster ("Other"; $n = 25$) contained a mixture of various taxa, including *Streptococcus agalactiae*, *Streptococcus anginosus*, *Prevotella*, and *Anaerococcus*, not all of which were detected in all samples within the cluster. Lastly, two samples characterized by high proportions of *Bifidobacterium* were grouped into a seventh cluster termed "Bifido"; this included a postmenarcheal sample from subject 124 and a sample from her mother. Vaginal pH varied significantly among clusters, with samples in groups GV and Other each having significantly higher pH than the LC and LI clusters (see Fig. S2 in the supplemental material).

Some clusters or community types appeared to be more commonly associated with either premenarcheal or postmenarcheal status in girls. Figure 2 compares the number and proportion of premenarcheal, postmenarcheal, and mother vaginal samples assigned to each cluster. The six clusters besides Bifido encompassed both premenarcheal and postmenarcheal vaginal samples from girls, but two clusters in particular were overrepresented in one group relative to the other. Among the premenarcheal samples ($n = 110$), 20.0% ($n = 22$) were assigned to the Other cluster, compared to only 3.4% ($n = 3$) of the postmenarcheal samples ($n = 87$), nearly a 6-fold difference. Those three postmenarcheal samples were all from subject 104, two of which had high proportions of *Streptococcus agalactiae* and one of which had a high proportion of *Peptoniphilus*. On the other hand, 50.6% ($n = 44$) of the postmenarcheal samples were assigned to cluster LC, compared to 22.7% ($n = 25$) of the premenarcheal samples, more than a 2-fold difference. Cluster GV was more common among mothers (25.5% of samples from mothers) than girls (9.1% collectively of pre- and postmenarcheal samples), but even so, finding *Gardnerella vaginalis* in the vaginal microbiota of girls prior to onset of partnered sexual activity is notable.

The hierarchical clustering results indicate that the vaginal microbiota of many premenarcheal girls were similar to those previously found in older adolescents and adults. To assess how long prior to menarche this was the case, we evaluated the cluster assignments over time within each individual (Fig. 3). Patterns of changes appeared to be highly individualized, but some common trends emerged. The vaginal microbiota of several girls were grouped into one of the *Lactobacillus*-dominated clusters (LC, LI, LG, or LJ) more than a year before menarche (e.g., subjects 101, 108, 127, and 128), many from the baseline visit. Some were initially in the Other cluster before transitioning to a *Lactobacillus* cluster (e.g., subjects 103, 111, 123, and 132). Except for subject 104, none transitioned back to the Other cluster following menarche. While some girls' microbiota remained in the same cluster over long periods of time (e.g., subjects 101 and 107), others transitioned between multiple groups (e.g., 103, 104, 112, and 126). We note that other transitions may have occurred during the intervals between quarterly visits, since evidence from studies of reproductive-age women indicates that short-term changes are commonplace (9, 10, 38–40). Nonetheless, communities typified by high proportions of *Lactobacillus* or other lactic acid bacteria (e.g., *Streptococcus*) were present well before the onset of menarche in most girls and were maintained after menarche.

Whereas hierarchical clustering was used to separate vaginal samples into major groups of community types, principal coordinates analysis (PCoA) was performed to obtain a more nuanced view of similarities and differences among samples in relation to other variables of interest. Figure S3A in the supplemental material shows that premenarcheal, postmenarcheal, and mother vaginal samples were similarly distributed, suggesting that differences in community composition are not strongly accounted for solely by age. Figure S3B, the same PCoA plot color-coded by hierarchical cluster assignment, emphasizes the variability within clusters and reveals that some groups are more distinct while others are spread across a broader range of space. This serves as an important reminder that bacterial composition and rank abundances within communities often lie along a continuum and need not be regarded as discrete types; rather, many communities may best be

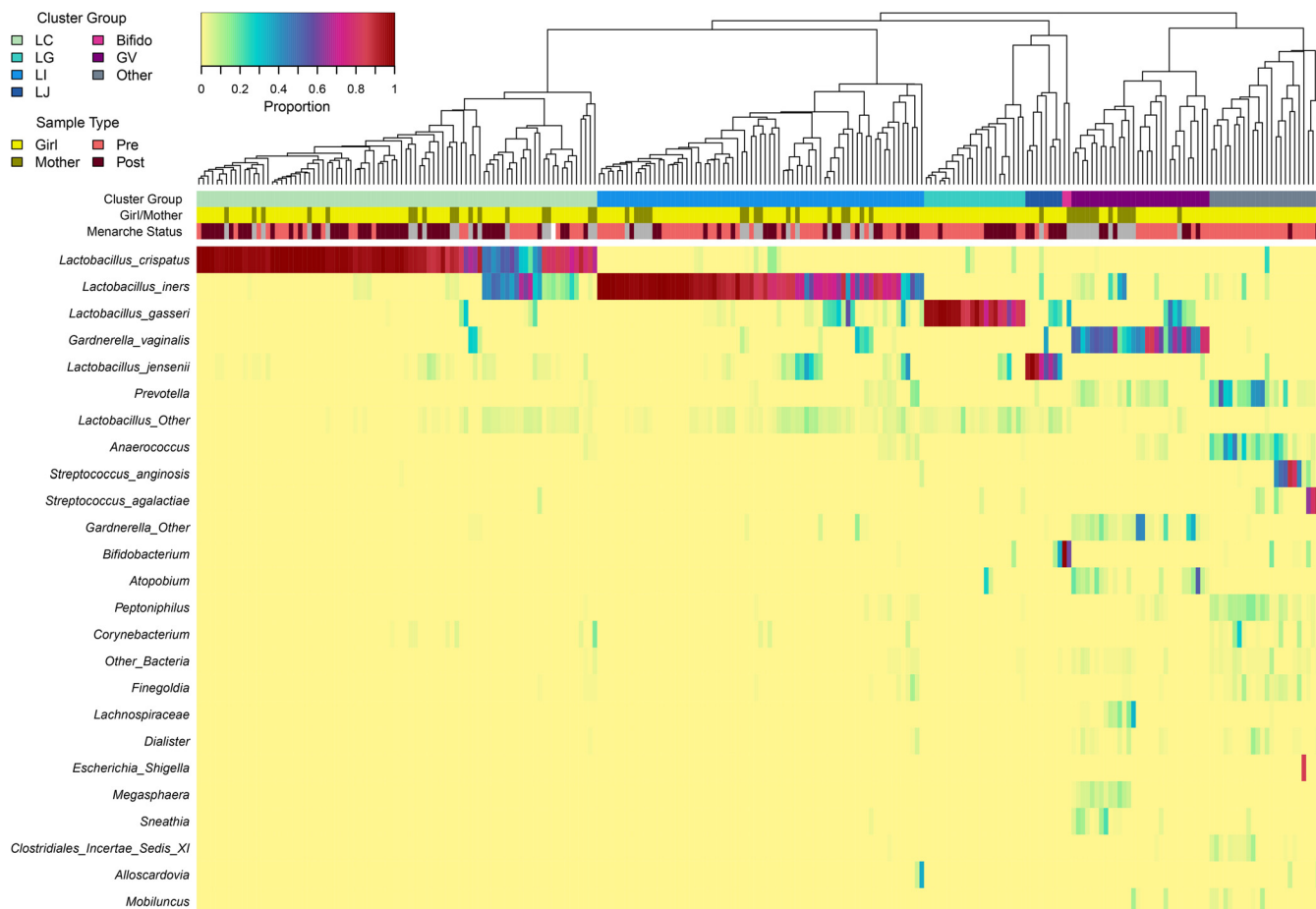


FIG 1 Bacterial community composition of the vaginal microbiota of girls and mothers sampled longitudinally. Each column in the dendrogram and heatmap represents the vaginal microbiota sampled from a single individual at a single time point. In total, 198 samples from 31 girls and 47 samples from 24 mothers are represented. The dendrogram represents the average linkage (UPGMA) hierarchical clustering of samples based on the Bray-Curtis dissimilarity matrix computed from Hellinger-standardized taxon abundance data. The colored bars below the dendrogram represent cluster group (top row) and sample type (second and third rows). Clusters are named to signify the most abundant taxon, when applicable: LC (*Lactobacillus crispatus* dominated; $n = 87$), LI (*L. iners*; $n = 71$), LG (*L. gasseri*; $n = 22$), LJ (*L. jensenii*; $n = 8$), Bifido (*Bifidobacterium*; $n = 2$), GV (*Gardnerella vaginalis*; $n = 30$), and Other ($n = 25$). The heatmap represents proportions (before Hellinger standardization) of the 25 overall most abundant taxa within each community, as indicated by the legend at top left. Sample type categories include girl/mother and premenarche/postmenarche (no menarche status is indicated for mother samples, in gray).

considered intermediate between groups characterized by different distributions of taxa.

Longitudinal dynamics of the perimenarcheal vaginal microbiota. To gain a detailed view of changes in community composition that occurred over time within each individual, we prepared summary plots of vaginal microbiota composition and associated metadata over time (Fig. 4). As anticipated based on the hierarchical clustering results, the vaginal microbiota of nearly all participants who reached menarche ($n = 21$) became characterized by a dominance of lactobacilli before or shortly after menarche. Interestingly, the microbiota of subjects 115 and 120 shifted from *Lactobacillus*-dominated in premenarche to *Gardnerella*-dominated in postmenarche. Even among premenarcheal vaginal microbiota, lactobacilli (primarily *Lactobacillus* spp., but in some cases *Streptococcus*) constituted at least 10% of the community in 27 girls (87.1%) and at least 50% in 25 girls (80.5%).

Figure 5 shows the progression of changes in the vaginal microbiota of four girls who had *Lactobacillus*-dominated communities before or at menarche, although the species composition

and temporal dynamics differed considerably. For instance, while *L. crispatus* dominated the microbiota of subject 107 for more than 2 years, the microbiota of subject 109 gradually shifted from *L. iners* dominated to *L. crispatus* dominated. Subject 102 had *Gardnerella vaginalis* and *L. gasseri* in premenarche but transitioned to *L. iners* following menarche. Subject 103 began the study 2 years before menarche with a diverse assortment of anaerobes that would usually be considered typical of prepubertal vaginal microbiota, but she eventually developed a microbiota dominated by *L. jensenii*, *L. gasseri*, and *Bifidobacterium*. These examples demonstrate several trajectories to *Lactobacillus*-dominated vaginal microbiota and emphasize the idea that the establishment of *Lactobacillus* dominance does not necessarily result in static community composition. Moreover, multiple species of *Lactobacillus* may be numerically dominant at different times in the same individual, consistent with observations in adult women (9). Community profiles for all participants, complete with associated vulvar and mothers' vaginal microbiota where applicable, are shown in Fig. S4 in the supplemental material.

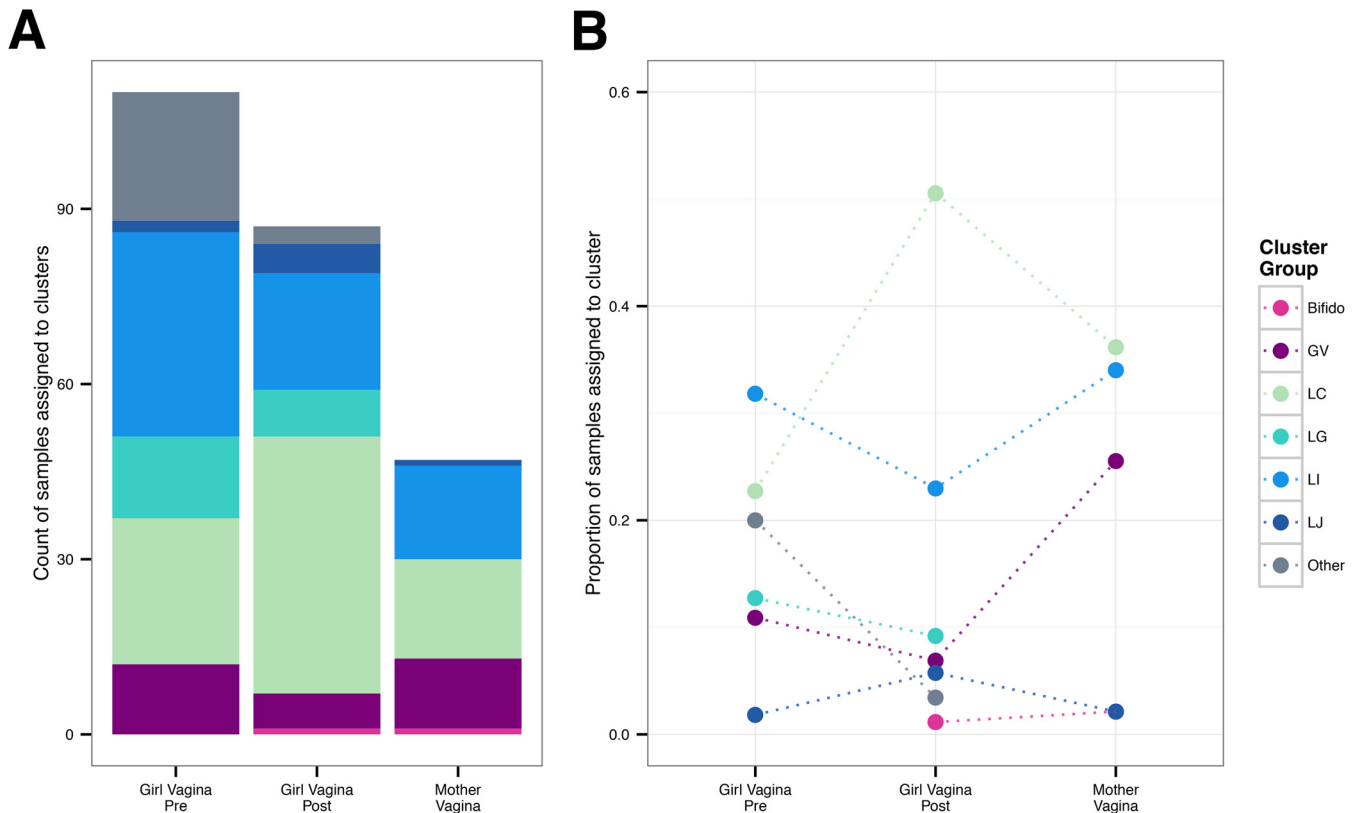


FIG 2 Hierarchical cluster assignment by sample type. One hundred ninety-eight vaginal microbiota from 31 girls and 47 vaginal microbiota from 24 mothers were separated into seven groups by hierarchical clustering. (A) Count of girl premenarcheal ($n = 110$), girl postmenarcheal ($n = 87$), and mother ($n = 47$) vaginal microbiota assigned to each cluster group (cluster names are the same as in Fig. 1). (B) Proportion of samples in each group assigned to each cluster. The dotted lines serve to highlight differences between sample types and do not represent changes in cluster group prevalence over time.

***Gardnerella vaginalis* in the perimenarcheal vaginal microbiota.** *Gardnerella vaginalis* was commonly detected among adolescent participants, constituting 10% or more of the vaginal microbiota in at least one sampling of 11 girls (35.5%), seven of whom had communities dominated by *Gardnerella* at some point in time. The proportion of *Gardnerella* in these participants over time is shown in Fig. S5 in the supplemental material. Some girls had high proportions of *Gardnerella* at multiple consecutive visits (e.g., subjects 102, 121, and 135), while it was detected in others at only one time point (e.g., subjects 112, 116, and 126). Participants with *Gardnerella* represented both black (7/21, 33.3%) and white (4/8, 50.0%) racial groups. Although *G. vaginalis* is commonly associated with bacterial vaginosis (BV) and some have suggested that it is a sexually transmitted infectious agent (41, 42), these findings present clear evidence that it can reside in the vaginal microbiota of healthy perimenarcheal girls without a history of partnered sexual behavior or symptoms characteristic of BV (both of which were exclusion criteria).

Lactic acid bacteria and vaginal pH. High numbers of lactobacilli, particularly species of *Lactobacillus*, are widely regarded as a hallmark of vaginal health in adult women. Their presence is typically associated with a low vaginal pH of around 4 to 4.5, although healthy women in black and Hispanic groups have been shown to have slightly higher pH on average (8). Since we observed high proportions of lactobacilli in the adolescent participants in this study, we next examined the relationship between the

proportions of lactobacilli and vaginal pH with respect to pubertal development and menarche. In this case, we defined lactic acid bacteria (LAB) as the genera in our data set contained within the order Lactobacillales, which included *Lactobacillus*, *Streptococcus*, *Aerococcus*, and *Facklamia*. We note that other genera, such as *Atopobium* and *Bifidobacterium*, are also known producers of lactic acid, so our representation of LAB is therefore somewhat conservative.

Overall, the relative abundance of LAB tended to increase with pubertal development, while vaginal pH tended to decrease (Fig. 5). We compared several linear mixed-effects models, controlling for intersubject variation, to determine how Tanner stage, age, and menarcheal status were related to changes in LAB proportions and pH. Because Tanner breast and pubic stage were collinear and accounted for similar proportions of variance (analysis at <https://github.com/roxanahickey/adolescent/blob/master/03-community-dynamics.md>), we tested models that included either breast or pubic scores, but not both. We focus on models using Tanner breast scores here (Table 2), and models with Tanner pubic scores are reported in Table S2 in the supplemental material. Following a stepwise model selection strategy, the optimal model accounting for changes in LAB proportions included only age and Tanner breast stage as fixed effects, since inclusion of menarche status added little explanatory value. Under this model, the transition from Tanner breast stage 2 to 3 was associated with a highly significant increase in logit-transformed LAB proportions

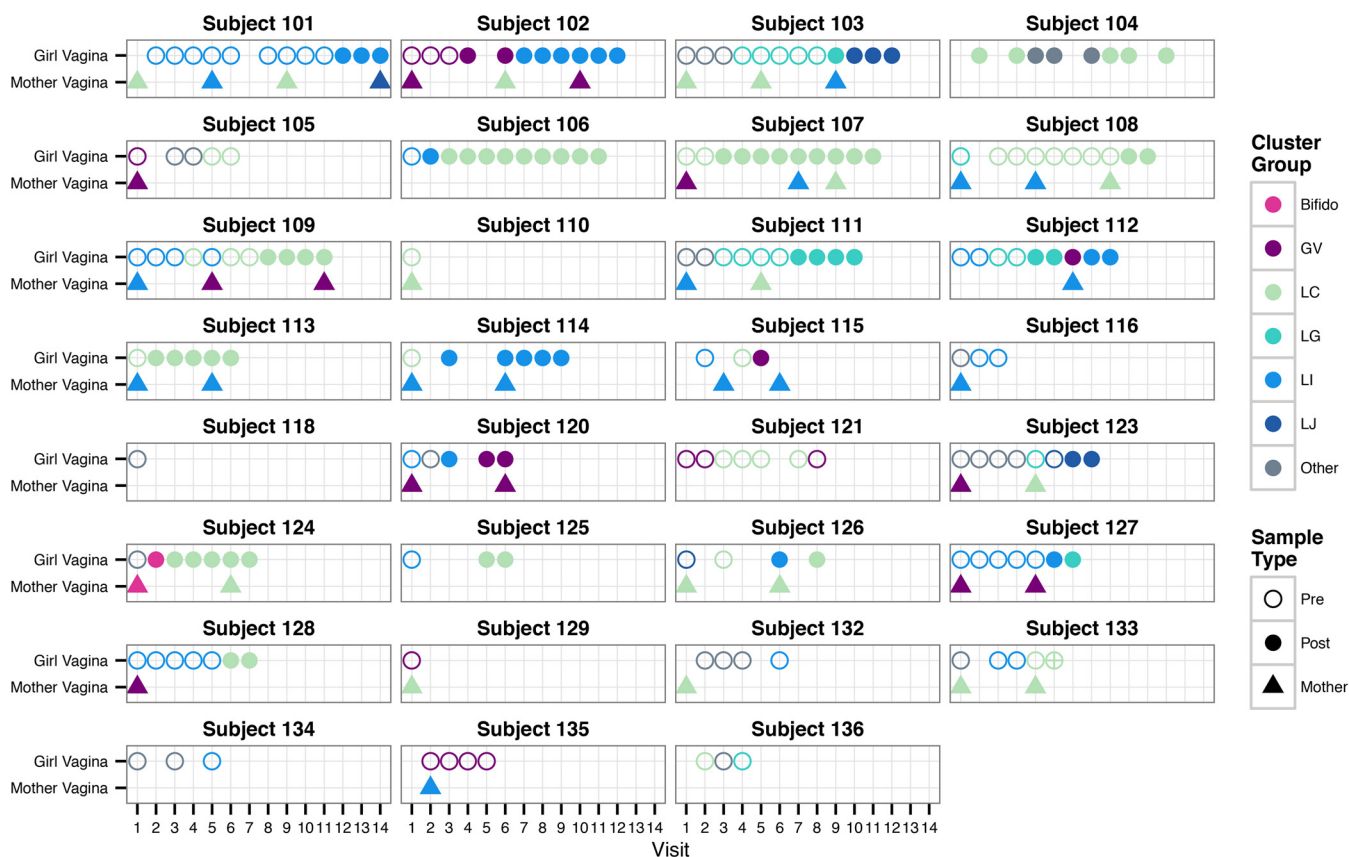


FIG 3 Hierarchical cluster assignment over time within individual participants. Each panel shows the hierarchical cluster assignment (same as in Fig. 1) of vaginal microbiota samples from an individual girl (circles) and her mother (triangles), when applicable. The x axis indicates the clinical visit at which each sample was collected (visits occurred approximately every 3 months). Open circles signify premenarcheal status, and filled circles signify postmenarcheal status in girls. The menarcheal status was not recorded for subject 133 at visit 6, indicated by an open circle with a cross.

($P = 1.1\text{E-}05$), while subsequent transitions to stage 4 and 5 had relatively small effects ($P = 0.26$ and $P = 0.56$, respectively [Table 2]). A similar pattern was detected using Tanner pubic scores instead (Table S2). The best model accounting for changes in vaginal pH included Tanner breast stage, menarche status, and age as fixed effects. Although the overall model accounted for a significant proportion of variance, only age had a significant marginal effect on pH ($P = 0.0071$), and the transitions in Tanner stage and menarche status were not themselves significantly associated with changes in pH (Table 2). However, a similar model using Tanner pubic scores indicated the transition from stage 2 to 3 was significantly associated with declining pH ($P = 0.0006$), while subsequent transitions were not significant (Table S2). Taken together, these results suggest the transition from Tanner breast or pubic stage 2 to 3 (i.e., early to midpuberty) is associated with the most profound changes in LAB proportions and vaginal pH.

Although increased LAB proportions were often coupled with decreased vaginal pH over time, closer inspection revealed that many vaginal samples with high proportions of LAB still had a pH well above 4.5, including both premenarcheal and postmenarcheal samples at all Tanner stages. We hypothesized this may be due to lower bacterial loads in the perimenarcheal vaginal microbiota, leading to lower levels of lactic acid and elevated pH. To test this hypothesis, we performed qPCR to estimate the number of

16S rRNA gene copies in 24 randomly selected girl vaginal samples with high proportions of LAB (>0.75) and either “low” (<5.0 , $n = 12$) or “high” vaginal pH (≥ 5.0 , $n = 12$). A detailed explanation of our approach and findings is summarized in Text S1 in the supplemental material. Although we did not find a statistically significant difference in the average estimated number of 16S rRNA gene copies in the “low” versus “high” pH groups ($P = 0.14$), we suggest that this phenomenon warrants consideration in future studies.

Vulvar microbiota of perimenarcheal girls. Although this study was primarily aimed at characterizing the vaginal microbiota of perimenarcheal girls, we also analyzed the composition of vulvar swabs and compared them to the vaginal microbiota. This included making qualitative comparisons, calculating correlations between paired vagina-vulva microbiota in taxonomic rank and relative abundances, performing hierarchical clustering as we did for the vaginal samples, and performing “indicator species” analysis (43, 44) to determine whether certain taxa were more closely associated with the vaginal or vulvar environment.

Composition of the vulvar microbiota typically mirrored that of contemporaneously sampled vaginal microbiota, although the vulva tended to have a greater variety of bacterial taxa. Hierarchical clustering analysis of vaginal and vulvar samples together resulted in the same groups overall (see Fig. S6 in the supplemental material), although the distribution of samples across groups was

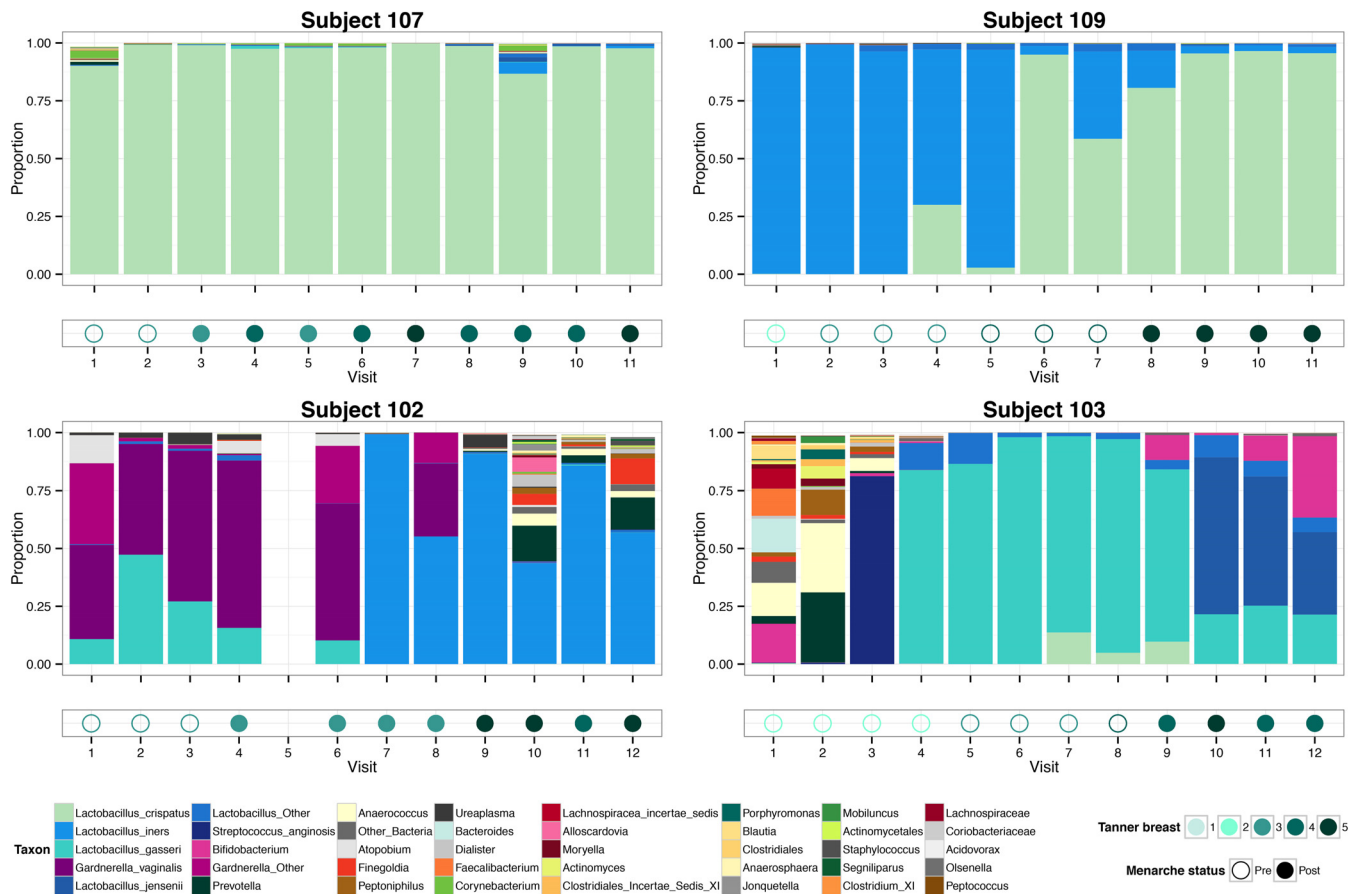


FIG 4 Transitions to *Lactobacillus*-dominant vaginal microbiota in four perimenarcheal girls. Panels show the vaginal bacterial community profiles and associated pubertal development of four participants sampled longitudinally. Bar plots represent the proportions of bacterial taxa in the community (legend at bottom left). Below each bar plot, the menarcheal status and clinician-assessed Tanner stage of breast development are indicated by symbol fill and color, respectively (legend at bottom right). Empty spaces in the plots indicate a skipped visit.

slightly different (e.g., the *L. jensenii* cluster was much smaller, with many samples assigned instead to the Other or *L. iners* cluster). Spearman's rank correlation coefficients were computed for the 186 matched vagina-vulva pairs based on the taxon proportion data. The mean correlation was 0.63 (the median was also 0.63), indicating a moderately high degree of concordance in the rank-order of taxa in paired vaginal and vulvar samples. Indicator species analysis, summarized in Table S3 in the supplemental material, revealed that three taxa were more strongly associated with the vulva (i.e., *Segniliparus*, *Murdochella*, and *Fusobacterium*), and each of these was typically minor in relative abundance (on average, 1.0 to 3.5% in the vulva compared to 0.02 to 0.04% in the vagina). No taxa were more closely associated with the vagina than the vulva, indicating that most vaginal species are likely also found in the vulvar microbiota. Seventeen taxa were associated jointly with the premenarcheal vagina, premenarcheal vulva, and postmenarcheal vulva, suggesting that, at least in this data set, the premenarcheal vaginal microbiota had more taxa in common with the vulvar microbiota than did the postmenarcheal vaginal microbiota. These findings suggest the vulvar microbiota of girls are generally similar to vaginal microbiota, perhaps even more so prior to menarche.

We qualitatively assessed changes in genus-level richness and Simpson's diversity index of vaginal and vulvar microbiota in re-

lation to Tanner stage and menarche status (see Text S1 in the supplemental material). As suggested by the indicator species analysis, the vulva tended to have greater richness and diversity than the vagina. Decreases in diversity with Tanner stage progression mirrored the increases in LAB proportions. However, lower richness and diversity in the vaginal microbiota were not necessarily associated with lower vaginal pH.

Similarities between vaginal microbiota of girls and their mothers. Lastly, we assessed similarities between the vaginal microbiota of girls and their mothers. This analysis was exploratory, as we had no basis on which to expect more or less similarity within girl-mother pairs. We qualitatively assessed similarities by studying the hierarchical cluster assignments presented in Fig. 3 and the paired community composition profiles depicted in Fig. S4 in the supplemental material. There was no general trend of resemblance between girls and mothers, but some interesting similarities were observed on a case-by-case basis. Some pairs often shared the same dominant taxon (e.g., *G. vaginalis* in subject pair 102/202, *L. iners* in 109/209 and 114/214, and *Bifidobacterium* or *L. crispatus* in 124/224), although even in those cases this similarity did not persist over repeated visits. In several cases, girls' microbiota bore virtually no resemblance to their mothers' (e.g., subject pairs 111/211, 113/213, and 132/232). The data are insufficient to determine whether any of these patterns are biologically meaning-

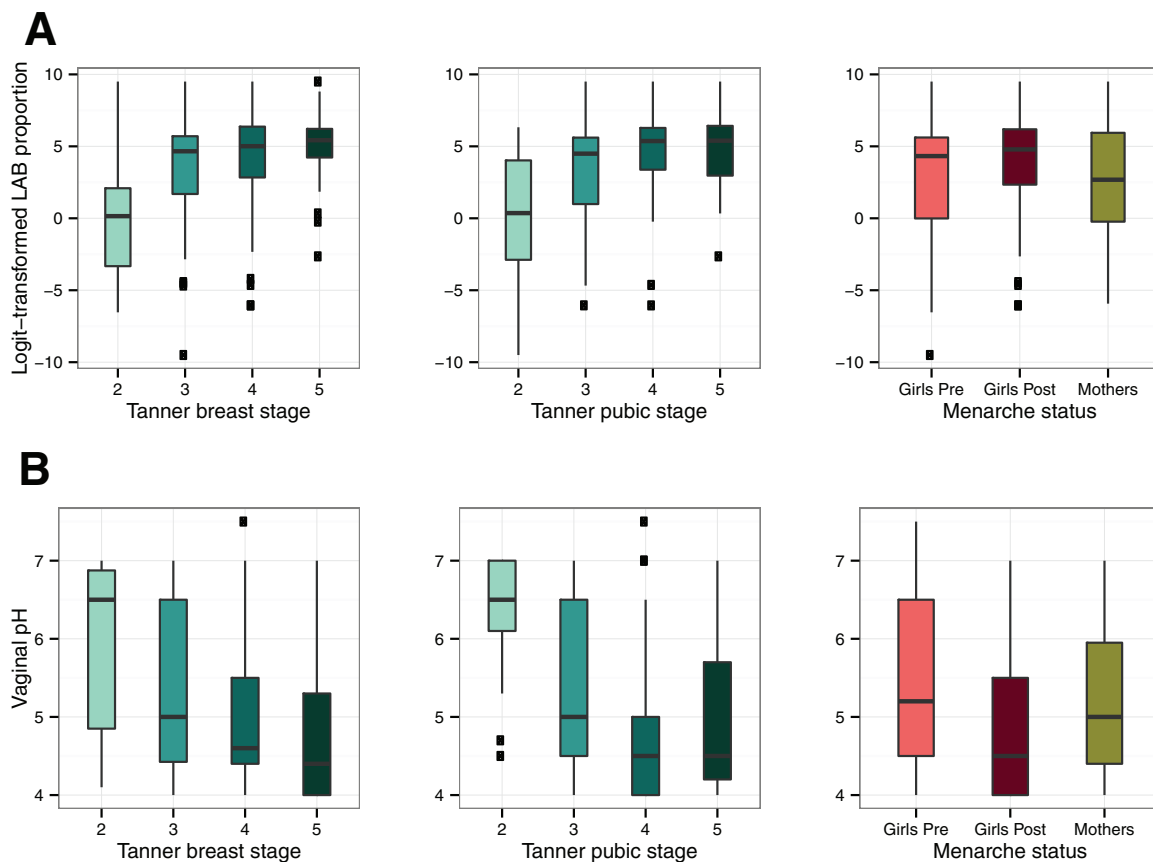


FIG 5 Trends in relative abundance of lactic acid bacteria and vaginal pH with pubertal development and menarche status. Box plots of the logit-transformed proportion of lactic acid bacteria (LAB; includes *Lactobacillus*, *Streptococcus*, *Aerococcus*, and *Facklamia*) (A) and vaginal pH of 31 perimenarcheal girls (B) are shown. Box plots show the relationship to Tanner breast stage (left), Tanner pubic stage (middle), and menarche status (right). The right column also includes data for 24 mothers who participated in the study. In each plot, the box represents the interquartile range, the whiskers represent the upper and lower quartiles, the horizontal line represents the median, and black square dots represent outliers.

ful, but the observation of similar vaginal microbiota in at least some girl-mother pairs is intriguing and warrants further studies aimed at understanding how the vaginal microbiota may be influenced by both genetic and environmental factors.

DISCUSSION

Vaginal microbiota play an important protective role in women's health, but our understanding of the microbiota in pre- and perimenarcheal girls has thus far been limited by a lack of detailed studies using modern techniques to assess community composition. In this prospective longitudinal study, we characterized changes in the vaginal and vulvar microbiota of 31 girls as they progressed through puberty to menarche. To our knowledge, it is the first study of perimenarcheal girls to employ high-throughput 16S rRNA gene sequencing, enabling finer resolution of the species and genus membership of communities, along with measurements of pubertal development (Tanner stages) and vaginal pH. Our findings add significant new insight to the composition and dynamics of the vaginal microbiota of adolescents during puberty and serve as a foundation for future studies aimed at characterizing the vaginal microbiota over shorter time intervals and for longer spans of time, opening the door to the possibility of identifying drivers of community composition.

The vaginal microbiota of pre- and perimenarcheal girls in our

study were similar to those previously described in older postmenarcheal adolescents (13, 14) and adults (7, 8, 45–47) inasmuch as lactic acid bacteria (LAB) were prominent members of most communities. For many girls, this was true well before menarche, often a year or more in advance. This was unanticipated given the results of previous studies that employed cultivation-dependent methods and found lactobacilli to be infrequent or minor constituents of the microbiota in premenarcheal girls (1, 3, 18, 48). However, those studies did not focus on the appearance of adult-like microbiota during pubertal transitions. In our study we tracked these transitions using Tanner's criteria for breast and pubic development (33), which are commonly used by clinicians to assess physical maturity. The five Tanner stages represent major changes in secondary sex characteristics that females undergo from prepuberty (stage 1) to full maturity (stage 5) in response to increased estrogen levels and other physiological changes. Menarche is considered a late event in puberty, usually occurring in stage 3 or 4 (34). We observed the most substantial increases in proportions of LAB, along with declines in vaginal pH, in the transition from Tanner breast stage 2 to 3, followed by less pronounced changes thereafter. Most of the microbiota with low proportions of or no LAB detected were obtained from girls earlier in puberty (breast stage 2 or 3 and premenarcheal). These communities were typically composed of an assortment of strict and facultative anaer-

TABLE 2 Linear mixed effects modeling of lactic acid bacteria and vaginal pH^a

Model and parameters	Result for model			
LAB model [logit(LAB) ~ TB + age + 1 subject + ε]				
Random effects	Variance	SD ^b	No. of observations	No. of groups
Subject (intercept)	4.4	2.1	189	28
Residual	7.3	2.7		
Fixed effects/contrasts ^b	Coefficient ^c	SE	df	P value ^d
Intercept	−6.1	4.4	150.6	1.7E-01
TB 3 vs. 2	3.3	0.7	183.9	1.1E-05***
TB 4 vs. 3	−0.6	0.5	176.2	2.6E-01
TB 5 vs. 4	0.4	0.7	176.7	5.6E-01
Age	0.7	0.4	151.7	4.1E-02*
Vaginal pH model (pH ~ TB + menarche status + age + 1 subject + ε)				
Random effects	Variance	SD	No. of observations	No. of groups
Subject (intercept)	0.6	0.8	122	20
Residual	0.4	0.6		
Fixed effects/contrasts ^b	Coefficient ^c	SE	df	P value ^d
Intercept	9.3	1.4	112.5	2.1E-09***
TB 3 vs. 2	−0.3	0.2	108.6	2.1E-01
TB 4 vs. 3	0.0	0.2	105.5	8.4E-01
TB 5 vs. 4	0.0	0.2	104.2	8.6E-01
Postmenarche vs. premenarche	−0.2	0.2	115.5	3.7E-01
Age	−0.3	0.1	113.4	7.1E-03**

^a LAB, lactic acid bacterium proportion; TB, Tanner breast stage; ε, random error; SD, standard deviation; SE, standard error; df, degrees of freedom.

^b Contrasts between successive Tanner stages were made, excluding stage 1 (not represented).

^c Marginal slope of the fixed effect on the response.

^d Significance is indicated as follows: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

obes, an observation more in line with what has been observed previously in premenarcheal girls.

Interestingly, high proportions of LAB in girls were not always associated with low vaginal pH (i.e., pH ≤ 4.5, but see reference 8), as is commonly observed in reproductive-age women with similar microbiota composition. We postulated that this could be due to lower bacterial loads (and therefore less lactic acid production) but were unable to detect a statistically significant difference in 16S rRNA gene copy number as a proxy for bacterial cell number between a subset of samples with high proportions of lactobacilli and either low or high pH. A study by Brabin et al. (21) found that higher vaginal pH occurred more often in adolescents with abnormal than normal menstrual cycles; this is another possible explanation for the patterns observed in our data. Our findings resonate with those reported by Thoma et al. that assessed longitudinal changes in the vaginal microbiota of 13- to 18-year-olds using Nugent Gram stain criteria (15). That study enrolled 49 virginal girls from a rural district of Uganda and obtained samples from them weekly for 2 years. Among premenarcheal girls the authors observed a significant increase in large Gram-positive rods (presumably *Lactobacillus* spp.) and a concurrent decrease in small Gram-negative to variable rods (presumably *G. vaginalis* or *Bacteroides* spp.) over time. In contrast, changes in the abundance of those morphotypes among the perimenarcheal and postmenarcheal groups were attenuated over time. Furthermore, while declines in vaginal pH over time were observed for all three groups, the trend was statistically significant only in the perimenarcheal and postmenarcheal groups. The authors concluded that vaginal microbiota transitioning occurs prior to menarche but that significant decreases in vaginal pH continue well after menarche. Similarly, we observed the steepest changes in relative abundance of lactobacilli in early to midpuberty, with less pronounced increases in later stages of puberty. The observation of Thoma et al. of ac-

celerated declines in vaginal pH after menarche (15) is also consistent with our finding that pH often remained higher than expected even though lactobacilli were numerically dominant.

Although the emergence of lactobacilli as dominant members of the microbiota in early puberty to midpuberty was a consistent trend among our participants, community composition and dynamics varied considerably between and within individuals. Various lactobacilli were dominant in different community types, including *Lactobacillus crispatus*, *L. iners*, *L. gasseri*, *L. jensenii*, and, in some cases, *Streptococcus* spp. The significance of communities dominated by different lactobacilli is not well understood, but recent studies have identified variation in the genetic and metabolic potential of vaginal lactobacilli that may influence the ecology of each species or strain *in vivo* (49, 50). Furthermore, some girls had only one dominant species at consecutive 3-month intervals, while others exhibited co-occurrence of two or more lactobacilli or a succession of multiple species over time. As previously seen in longitudinal studies of reproductive-age women, transitioning through multiple community state types seems to be at least as common as having a community with stable or constant species composition over time (9, 10). Our findings suggest that this pattern extends to earlier stages of life as well. This is an important reminder that the vaginal ecosystem is highly dynamic and individualized, and both temporal and interindividual differences should be taken under consideration as we attempt to characterize and understand the “normal” vaginal microbiota. Larger, longitudinal studies extending into late adolescence will be required to investigate the relevance of dynamic vaginal microbiota during puberty to long-term vaginal health.

The high prevalence of *Gardnerella vaginalis* among adolescent participants was a surprising finding in this study. It was the fourth most abundant species detected and constituted at least 10% of the vaginal microbiota at least once in approximately one-

third of the participants. In some cases, *G. vaginalis* constituted the majority of the microbiota at multiple consecutive 3-month intervals, suggesting long-term persistence. *G. vaginalis* is often associated with bacterial vaginosis (BV), a common yet poorly understood condition associated with increased risk of sexually transmitted infections and preterm birth (51, 52). While *G. vaginalis* is undoubtedly correlated with BV, evidence for whether it is necessary or sufficient to elicit symptoms remains inconclusive (42, 53–55). Postulated virulence mechanisms include biofilm formation (56), sialidase production (57), and synergism with other bacteria, such as *Prevotella* (58). However, numerous studies have identified *G. vaginalis* in many reportedly healthy women, suggesting that it may be a commensal in some individuals (7, 8, 46, 59). Considering evidence that identical *G. vaginalis* 16S rRNA sequences are detected among sexual partners (60), some have suggested that *G. vaginalis* is acquired exclusively through sexual contact (42). Our findings present evidence contrary to this hypothesis, as none of the girls in this study had a history of partnered sexual activity. Likewise, several studies report finding *G. vaginalis* in children and adolescents without sexual experience (18, 19, 48, 61–63). Whether microbiota containing *G. vaginalis* place the host at higher risk of infection remains to be seen, but recent investigations into the remarkable diversity of this species (64–67) may help to clarify whether commensal and pathogenic strains exist and can be distinguished from one another.

Study of the vaginal microbiota in adolescents may provide crucial insight into the complexity and variability of microbiota in adults, as well as whether the development of microbial communities during puberty is associated with vaginal health in adulthood. Our research demonstrates the feasibility of longitudinal study of girls to complement understanding of the vaginal microbiome of adult women. The 10- to 12-year-old participants reported feeling comfortable and safe with the study procedures, especially after the initial visit (68). The small, flexible swabs easily passed patent hymens and captured sufficient vaginal material for analysis. A trained female provider and accompaniment by the subjects' mothers also allowed girls to become rapidly accustomed to sample acquisition and complete most scheduled study visits. Only six girls declined vaginal samples during the baseline visit, although our results, and those of others, suggest that vulvar samples may serve as a rough proxy for vaginal samples (69–71) if vaginal samples are declined. Similar studies can and should be done to grow our understanding of both normal and abnormal vaginal microbiota and ultimately improve strategies of gynecologic care for adolescent girls.

Many questions remain unanswered, including those regarding the relationship between estrogen, vaginal glycogen levels, lactic acid bacteria abundance, and vaginal pH; factors that determine successional changes at earlier stages of puberty; and whether differences in community composition during adolescence can influence health outcomes later in life. Although the concordance between girls and mothers was inconsistent, future studies of girl-mother or sister-sister pairs could elucidate the role of host genetics and interactions among individuals in shaping the microbiota. To gain a comprehensive understanding of the importance of vaginal microbiota composition and dynamics throughout puberty, it would be wise to include girls on the brink of puberty and even earlier in childhood, as well as older girls experiencing regular menstrual cycling. Ideally, similar studies should be performed on larger cohorts of girls with a balanced

sampling of racial and ethnic groups to determine whether the patterns we observed are more broadly conserved.

MATERIALS AND METHODS

Study design and enrollment criteria. The study protocol was approved by the Institutional Review Board of Indiana University, and informed consent from girls and their mothers was documented before participation in the study. Participants and their mothers were recruited by referral from clinicians, by referral from participating mothers, or in response to advertisements placed in local newsletters and newspapers. Primary inclusion criteria included an age of 10.0 to 12.9 years and premenarcheal status at enrollment, and initiation of pubertal development indicated by breast development of Tanner stage 2 or 3 (33), as documented by self-reporting and corroboration by a clinician's examination. Eligible candidates were enrolled if they were in good health and willing to refrain from bubble baths and genital cleansing wipes for 48 h before examination. Exclusion criteria included genitourinary symptoms (dysuria, vaginal discharge, or genital ulcers); evidence of urinary tract infection; use of any systemic or topical antibiotic or antifungal treatment within the previous 60 days; prior genitourinary surgery, instrumentation, or medical treatment for recurrent urinary tract infection, posterior urethral valves, or enuresis; any significant medical condition deemed cause for exclusion by the investigator (e.g., type I diabetes, asthma, or autoimmune diseases); prior evaluation for vulvovaginal symptoms; history of prepubertal bleeding; history of sexual abuse or sexual activity; and having already begun menarche.

Mothers of enrolled girls were invited to participate in annual sample collections. No age range was required, nor was biological maternity. Mothers were included if they reported themselves in good health and were willing to refrain from using bubble bath, douching substances, powders, perfumes, wet wipes, or lotions to the genital area for 48 h before sample collection and were willing to refrain from sexual activity, bathing, or swimming for 2 h before sample collection. Lack of participation by the mother was not an exclusion factor for otherwise eligible girls.

The study initially enrolled 32 girls and 25 mothers, but one participant (subject 117) was withdrawn at the baseline visit after the clinician determined she was at Tanner stage 1 for breast development, and no samples were collected. A sample collected from her mother (subject 217) was subsequently excluded from further analysis. Two participants (subjects 101 and 129) were also enrolled at breast stage 1 and permitted to continue at the investigator's discretion, but subject 129 was lost to observation after the baseline visit. One participant (subject 110) was enrolled at breast stage 5 (her self-assessment was stage 3) but then lost to observation after two visits. Vaginal and vulvar samples from the latter three participants, along with any vaginal samples from their mothers, were retained for analysis.

Collection of specimens and participant metadata. At enrollment (baseline) and each quarterly visit, a female clinician completed a short physical examination to document breast, pubic hair, and genital development. Girls were then assisted to a sitting, "frog-legged" position with their feet in stirrups. After additional inspection of the vulva, the clinician sequentially obtained two vulvar samples by rubbing both labia minora with a sterile, dry, flocked nasopharyngeal swab. The clinician (after confirmation of the adequacy of the hymenal opening) then sequentially inserted up to three swabs through the vaginal introitus to approximately 5 cm. Vaginal swabs were rotated twice. Lastly, vaginal pH was obtained by inserting a commercial pH paper into the vaginal opening and noting the pH in comparison to a provided color indicator. Clinicians omitted pH measure at their discretion to optimize sample integrity. Study-provided cell phones were used to identify the onset of menses, and visits were scheduled so as not to coincide with menses. Participants' mothers provided up to three self-obtained vaginal specimens for each annual collection. Vulvar swabs were not collected from the mothers. All swab samples were placed separately in labeled cryovials that were immediately placed on dry ice and then transferred to and stored in a -70°C freezer.

Of the swabs collected from girls and mothers, one was shipped to the University of Idaho for analysis of microbial community composition as described below, and a second was archived for use in subsequent studies. The third swab was used to assess vaginal health by Nugent criteria (72). Vaginal or vulvar sample collection was halted if the patient wished to discontinue sampling during the examination, or at the clinician's discretion to optimize sample integrity.

16S rRNA sequencing and taxonomic assignment of reads. Genomic DNA was extracted from vaginal and vulvar swabs with the use of a validated enzymatic lysis and bead-beating protocol (73), followed by purification with use of the QIAamp DNA minikit (Qiagen, Venlo, Netherlands), which we have used in our previous human microbiome studies (8–10, 74). Bacterial 16S rRNA genes were amplified by PCR using bar-coded primers flanking hypervariable regions V1 and V3 (*Escherichia coli* positions 27F to 534R), optimized by Frank et al. (75) for improved detection of *Bifidobacteriaceae* (including *Gardnerella*), *Borrelia*, and *Chlamydia*, as done previously (10). Amplicons were sequenced on a Roche 454 FLX pyrosequencer (Roche 454 Life Sciences, Branford, CT, USA) at the University of Idaho. Sequence reads were cleaned, filtered, and taxonomically assigned using the Ribosomal Database Project (RDP) Naïve Bayesian Classifier to the first RDP level with a bootstrap score of ≥ 50 . Species of *Lactobacillus*, *Gardnerella*, and *Streptococcus* were further classified using a clustering approach with the R package WGCNA (76) and 16S rRNA sequences from the PATRIC database (<http://patricbrc.org>). The methods listed above are described in detail in Text S1 in the supplemental material. Following preprocessing of Roche 454 sequence data and taxonomic assignment of reads, data were processed using custom R scripts to calculate percentages of taxa within each sample. To simplify analysis of community composition, we retained named taxa that constituted either at least 1% of the community in two or more samples or at least 5% of the community in at least one sample. Taxonomically assigned reads that did not meet this threshold were combined into an “Other Bacteria” category, along with reads that could not be taxonomically assigned beyond the level of bacteria. Note that “Other Bacteria” is distinct from the “Other” group of vaginal and vulvar samples determined by hierarchical clustering.

Availability of data and custom R scripts. Sequences in standard flowgram format (SFF) for the 457 samples analyzed in this study are available for download from the NCBI Sequence Read Archive (BioProject PRJNA266340, <http://www.ncbi.nlm.nih.gov/bioproject/266340>). Data and custom R scripts to reproduce the analyses, including walkthroughs of intermediate steps as well as some additional analyses, are available on GitHub at <https://github.com/roxanahickey/adolescent>. Analyses were conducted using R v3.1.0 (35).

Hierarchical clustering and principal coordinates analysis of community composition. Following approaches outlined by Legendre and Legendre (77) and demonstrated by Borcard et al. (78), we performed both hierarchical clustering and PCoA to obtain an overall picture of similarities and differences in bacterial community composition across vaginal samples from girls and mothers. A similar analysis was performed for all vaginal and vulvar samples together. Prior to these analyses, the taxon abundance matrix was standardized using the Hellinger method with the R package *vegan* v2.0-10 (79). This approach is recommended when applying clustering or ordination techniques to species abundance data with sparse representation of some taxa among samples (80, 81). The Bray-Curtis dissimilarity (82) was then calculated from the Hellinger-transformed data and used to perform hierarchical clustering. Clustering was carried out using four different linkage methods: single, complete, average (unweighted pair group method with arithmetic mean [UPGMA]), and Ward's minimum variance criterion. The best clustering method was identified by determining the cophenetic distance (83) of each hierarchical clustering, followed by calculation of the Gower distance (36), the sum of squared differences between the original and cophenetic distances. The method with the smallest Gower distance was selected as the optimal clustering model for the distance matrix used. The optimum

number of clusters was selected according to the maximum silhouette width (37), and resulting cluster assignments were used in subsequent analyses as a categorical representation of community composition. Combined heatmap and dendrogram plots were generated using custom code that used the R package *gplots* v2.14.1 (84). A Kruskal-Wallis rank sum test was performed to determine whether vaginal pH differed significantly across hierarchical cluster groups (excluding Bifido due to having only two samples). This was followed by a pairwise multiple-comparison test using Tukey's method to identify significant differences between pairs of cluster groups.

As with hierarchical clustering, PCoA was performed with the Bray-Curtis dissimilarity matrix calculated from Hellinger-standardized taxon abundance data. To adjust for negative eigenvalues (the result of using a non-Euclidean distance matrix), a Cailliez correction was applied to the eigenvalues (85) before calculation of R^2 -like ratios (essentially, variance accounted for by each PCoA axis). PCoA plots were generated using the first two PCoA axes. These plots were used to make qualitative assessments of patterns in the participant metadata (e.g., sample type, Tanner stage, and menarcheal status) associated with differences in community composition.

Analysis of longitudinal trends in community composition and vaginal pH. Qualitative assessments of changes in community composition associated with pubertal development and menarcheal status were complemented by linear mixed effects modeling to identify variables significantly associated with observed patterns. We used the R package *lme4* v1.1-7 (86) to perform separate analyses of the relationships between lactic acid bacterium proportions or vaginal pH and selected participant metadata. LAB proportions were normalized by logit transformation prior to analysis (87). Subject was specified as a random effect in all models to control for interindividual variability, and we elected to exclude subjects with less than three observations of the response variable of interest to minimize inflation of error estimates. Fixed (i.e., explanatory) effects included Tanner breast or pubic stage (ordered factor with levels 1 through 5), menarche status (pre- or post-), and age at sampling. When Tanner stage was included in the model, we performed contrasts between progressive stages (e.g., stage 3 versus stage 2) to determine whether the response variable was significantly different between them. Starting with a simple model including Tanner stage as the sole fixed effect (e.g., $\text{LAB.logit} \sim \text{Tanner} + 1 | \text{Subject} + \epsilon$, where “1|Subject” specifies subject as a random effect and ϵ represents random error), we conducted a stepwise model comparison approach adding fixed effects and interactions one by one, using analysis of variance (ANOVA) to compare models at each step. If the models were not significantly different at an α value of 0.05, the simpler model was favored over the more complex model. *P* values for individual mixed effects models were obtained using the R package *lmerTest* v2.0-11 (88), which employs type III and type I *F* tests for fixed effects and likelihood ratio tests for random effects. Upon selection of the best model, residual and quantile-quantile plots were visually inspected to identify any obvious deviations from homoscedasticity or normality.

Comparisons of the vaginal and vulvar microbiota of girls. In addition to the hierarchical clustering and PCoA described above, indicator species analysis (44) was performed to identify bacterial taxa most strongly associated with groups of vaginal and vulvar samples. This was done by calculating indicator values with the *IndVal* function (89) in the R package *indicspecies* (43). This statistic represents the association of each taxon with one or more groups of samples. Groups were evaluated based on sample type (girl vagina, girl vulva, or mother vagina) and menarche status. Similarity between paired vaginal and vulvar samples from girls (i.e., samples collected at the same visit from the same individual) was assessed by calculating Spearman's rank correlation coefficient from taxon relative abundances.

Community richness and diversity analyses. We made qualitative assessments of changes in genus-level community richness and diversity (Simpson's index) in relation to pubertal development (i.e., Tanner stage and menarche status) and vaginal pH, detailed in Text S1 in the supple-

mental material. Rarefaction curves were used to determine a sampling depth of 2,000 observations per sample. Any taxa that could not be classified to the genus level were characterized as Other; therefore, the diversity estimates are somewhat conservative. Because of this, we elected not to perform additional quantitative analyses on these data.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00097-15/-/DCSupplemental>.

Text S1, PDF file, 0.2 MB.
Figure S1, PDF file, 0.01 MB.
Figure S2, PDF file, 0.04 MB.
Figure S3, PDF file, 0.03 MB.
Figure S4, PDF file, 0.2 MB.
Figure S5, PDF file, 0.01 MB.
Figure S6, PDF file, 0.1 MB.
Table S1, XLSX file, 0.04 MB.
Table S2, XLSX file, 0.04 MB.
Table S3, XLSX file, 0.03 MB.

ACKNOWLEDGMENTS

Financial support was provided by the Procter & Gamble Company and a grant to L.J.F. from the National Institute of General Medical Sciences of the National Institutes of Health (P30GM103324). R.J.H. was supported by the Bioinformatics & Computational Biology Graduate Program in partnership with IBEST.

We are grateful to Maria Schneider, Dorah Mtui, and Daniel Beck (University of Idaho) for their assistance in preparing genomic DNA samples and amplicons; staff of the Institute for Bioinformatics and Evolutionary Studies (IBEST) Genomics Resources Core and Computational Resources Core for technical support; Matthew Pennell (University of Idaho) for consultation on statistical analyses and reproducible research strategies; and James Foster (University of Idaho) and José Ponciano (University of Florida) for insightful discussions regarding data analysis.

REFERENCES

- Myhre AK, Bevanger LS, Berntzen K, Bratlid D. 2002. Anogenital bacteriology in non-abused preschool children: a descriptive study of the aerobic genital flora and the isolation of anogenital *Gardnerella vaginalis*. *Acta Paediatr* 91:885–891. <http://dx.doi.org/10.1111/j.1651-2227.2002.tb02850.x>.
- Hill GB, St Claire KK, Gutman LT. 1995. Anaerobes predominate among the vaginal microflora of prepubertal girls. *Clin Infect Dis* 20(Suppl 2): S269–S270. http://dx.doi.org/10.1093/clinids/20.Supplement_2.S269.
- Gerstner GJ, Grünberger W, Boschitsch E, Rotter M. 1982. Vaginal organisms in prepubertal children with and without vulvovaginitis. *Arch Gynecol* 231:247–252. <http://dx.doi.org/10.1007/BF02110125>.
- Hammerschlag MR, Alpert S, Onderdonk AB, Thurston P, Drude E, McCormack WM, Bartlett JG. 1978. Anaerobic microflora of the vagina in children. *Am J Obstet Gynecol* 131:853–856.
- Boskey ER, Telsch KM, Whaley KJ, Moench TR, Cone RA. 1999. Acid production by vaginal flora in vitro is consistent with the rate and extent of vaginal acidification. *Infect Immun* 67:5170–5175.
- Linhares IM, Summers PR, Larsen B, Giraldo PC, Witkin SS. 2011. Contemporary perspectives on vaginal pH and lactobacilli. *Am J Obstet Gynecol* 204:120.e1–120.e5. <http://dx.doi.org/10.1016/j.ajog.2010.07.010>.
- Zhou X, Brown CJ, Abdo Z, Davis CC, Hansmann MA, Joyce P, Foster JA, Forney LJ. 2007. Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women. *ISME J* 1:121–133. <http://dx.doi.org/10.1038/ismej.2007.12>.
- Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, Brotman RM, Davis CC, Ault K, Peralta L, Forney LJ. 2011. Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A* 108(Suppl 1):4680–4687. <http://dx.doi.org/10.1073/pnas.1002611107>.
- Gajer P, Brotman RM, Bai G, Sakamoto J, Schütte UM, Zhong X, Koenig SS, Fu L, Ma ZS, Zhou X, Abdo Z, Forney LJ, Ravel J. 2012. Temporal dynamics of the human vaginal microbiota. *Sci Transl Med* 4:132ra52. <http://dx.doi.org/10.1126/scitranslmed.3003605>.
- Ravel J, Brotman RM, Gajer P, Ma B, Nandy M, Fadrosch DW, Sakamoto J, Koenig SS, Fu L, Zhou X, Hickey RJ, Schwebke JR, Forney LJ. 2013. Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. *Microbiome* 1:29. <http://dx.doi.org/10.1186/2049-2618-1-29>.
- Verhelst R, Verstraeten H, Claeys G, Verschraegen G, Van Simaey L, De Ganck C, De Backer E, Temmerman M, Vaneechoutte M. 2005. Comparison between Gram stain and culture for the characterization of vaginal microflora: definition of a distinct grade that resembles grade I microflora and revised categorization of grade I microflora. *BMC Microbiol* 5:61. <http://dx.doi.org/10.1186/1471-2180-5-61>.
- Srinivasan S, Liu C, Mitchell CM, Fiedler TL, Thomas KK, Agnew KJ, Marrazzo JM, Fredricks DN. 2010. Temporal variability of human vaginal bacteria and relationship with bacterial vaginosis. *PLoS One* 5:e10197. <http://dx.doi.org/10.1371/journal.pone.0010197>.
- Alvarez-Olmos MI, Barousse MM, Rajan L, Van Der Pol BJ, Fortenberry D, Orr D, Fidel PL, Jr. 2004. Vaginal lactobacilli in adolescents. *Sex Transm Dis* 31:393–400. <http://dx.doi.org/10.1097/01.OLQ.0000130454.83883.E9>.
- Yamamoto T, Zhou X, Williams CJ, Hochwalt A, Forney LJ. 2009. Bacterial populations in the vaginas of healthy adolescent women. *J Pediatr Adolesc Gynecol* 22:11–18. <http://dx.doi.org/10.1016/j.jpog.2008.01.073>.
- Thoma ME, Gray RH, Kiwanuka N, Aluma S, Wang M-C, Sewankambo N, Wawer MJ. 2011. Longitudinal changes in vaginal microbiota composition assessed by Gram stain among never sexually active pre- and post-menarcheal adolescents in Rakai, Uganda. *J Pediatr Adolesc Gynecol* 24: 42–47. <http://dx.doi.org/10.1016/j.jpog.2010.07.002>.
- Joishy M, Ashtekar CS, Jain A, Gonsalves R. 2005. Do we need to treat vulvovaginitis in prepubertal girls? *BMJ* 330:186–188. <http://dx.doi.org/10.1136/bmj.330.7484.186>.
- Randjelović G, Kocić B, Stojanović M, Mišić M. 2005. Bacteriological findings of the vulvar swab specimens from girls with vulvovaginitis. *Facta Univ Med Biol* 12:159–163.
- Yilmaz AE, Celik N, Soyulu G, Donmez A, Yuksel C. 2012. Comparison of clinical and microbiological features of vulvovaginitis in prepubertal and pubertal girls. *J Formos Med Assoc* 111:392–396. <http://dx.doi.org/10.1016/j.jfma.2011.05.013>.
- Bump RC, Sachs LA, Buesching WJ. 1986. Sexually transmissible infectious agents in sexually active and virginal asymptomatic adolescent girls. *Pediatrics* 77:488–494.
- Schwebke JR, Desmond RA, Oh MK. 2004. Predictors of bacterial vaginosis in adolescent women who douche. *Sex Transm Dis* 31:433–436. <http://dx.doi.org/10.1097/01.OLQ.0000129948.91055.9F>.
- Brabin L, Roberts SA, Fairbrother E, Mandal D, Higgins SP, Chandiok S, Wood P, Barnard G, Kitchener HC. 2005. Factors affecting vaginal pH levels among female adolescents attending genitourinary medicine clinics. *Sex Transm Infect* 81:483–487. <http://dx.doi.org/10.1136/sti.2005.014621>.
- Brotman RM, Erbeling EJ, Jamshidi RM, Klebanoff MA, Zenilman JM, Ghanem KG. 2007. Findings associated with recurrence of bacterial vaginosis among adolescents attending sexually transmitted diseases clinics. *J Pediatr Adolesc Gynecol* 20:225–231. <http://dx.doi.org/10.1016/j.jpog.2006.11.009>.
- Schellenberg J, Blake Ball T, Lane M, Cheang M, Plummer F. 2008. Flow cytometric quantification of bacteria in vaginal swab samples self-collected by adolescents attending a gynecology clinic. *J Microbiol Methods* 73:216–226. <http://dx.doi.org/10.1016/j.mimet.2008.03.004>.
- Vaca M, Guadalupe I, Erazo S, Tinizaray K, Chico M, Cooper P, Hay P. 2010. High prevalence of bacterial vaginosis in adolescent girls in a tropical area of Ecuador. *BJOG* 117:225–228. <http://dx.doi.org/10.1111/j.1471-0528.2009.02397.x>.
- Merchant JS, Oh K, Klerman LV. 1999. Douching: a problem for adolescent girls and young women. *Arch Pediatr Adolesc Med* 153:834–837. <http://dx.doi.org/10.1001/archpedi.153.8.834>.
- Simpson T, Merchant J, Grimley DM, Oh MK. 2004. Vaginal douching among adolescent and young women: more challenges than progress. *J Pediatr Adolesc Gynecol* 17:249–255. <http://dx.doi.org/10.1016/j.jpog.2004.05.008>.
- Blythe MJ, Fortenberry JD, Orr DP. 2003. Douching behaviors reported by adolescent and young adult women at high risk for sexually transmitted

- infections. *J Pediatr Adolesc Gynecol* 16:95–100. [http://dx.doi.org/10.1016/S1083-3188\(03\)00027-5](http://dx.doi.org/10.1016/S1083-3188(03)00027-5).
28. Chase DJ, Schenkel BP, Fahr A-M, Eigner U, Tampon Study Group. 2007. A prospective, randomized, double-blind study of vaginal microflora and epithelium in women using a tampon with an apertured film cover compared with those in women using a commercial tampon with a cover of nonwoven fleece. *J Clin Microbiol* 45:1219–1224. <http://dx.doi.org/10.1128/JCM.02156-06>.
 29. Brotman RM, Klebanoff MA, Nansel TR, Andrews WW, Schwebke JR, Zhang J, Yu KF, Zenilman JM, Scharfstein DO. 2008. A longitudinal study of vaginal douching and bacterial vaginosis—a marginal structural modeling analysis. *Am J Epidemiol* 168:188–196. <http://dx.doi.org/10.1093/aje/kwn103>.
 30. Ott MA, Ofner S, Fortenberry JD. 2009. Beyond douching: use of feminine hygiene products and STI risk among young women. *J Sex Med* 6:1335–1340. <http://dx.doi.org/10.1111/j.1743-6109.2008.01152.x>.
 31. Karlsson CL, Molin G, Cilio CM, Ahrné S. 2011. The pioneer gut microbiota in human neonates vaginally born at term—a pilot study. *Pediatr Res* 70:282–286. <http://dx.doi.org/10.1203/PDR.0b013e318225f765>.
 32. Mirmonsef P, Gilbert D, Zariffard MR, Hamaker BR, Kaur A, Landay AL, Spear GT. 2011. The effects of commensal bacteria on innate immune responses in the female genital tract. *Am J Reprod Immunol* 65:190–195. <http://dx.doi.org/10.1111/j.1600-0897.2010.00943.x>.
 33. Tanner JM. 1962. Growth at adolescence, 2nd ed. Thomas, Springfield, IL.
 34. Marshall WA, Tanner JM. 1969. Variations in pattern of pubertal changes in girls. *Arch Dis Child* 44:291–303. <http://dx.doi.org/10.1136/adc.44.235.291>.
 35. R Core Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
 36. Gower JC. 1983. Comparing classifications, p 137–155. In Felsenstein J (ed), *Numerical taxonomy*. Springer, New York, NY.
 37. Kaufman L, Rousseeuw PJ. 2009. Finding groups in data: an introduction to cluster analysis. John Wiley & Sons, New York, NY.
 38. Santiago GL, Cools P, Verstraelen H, Trog M, Missine G, El Aila N, Verhelst R, Tency I, Claeys G, Temmerman M, Vaneechoutte M. 2011. Longitudinal study of the dynamics of vaginal microflora during two consecutive menstrual cycles. *PLoS One* 6:e28180. <http://dx.doi.org/10.1371/journal.pone.0028180>.
 39. Lopes dos Santos SG, Tency I, Verstraelen H, Verhelst R, Trog M, Temmerman M, Vancoillie L, Decat E, Cools P, Vaneechoutte M. 2012. Longitudinal qPCR study of the dynamics of *L. crispatus*, *L. iners*, *A. vaginae*, (sialidase positive) *G. vaginalis*, and *P. bivia* in the vagina. *PLoS One* 7:e45281. <http://dx.doi.org/10.1371/journal.pone.0045281>.
 40. Lambert JA, John S, Sobel JD, Akins RA. 2013. Longitudinal analysis of vaginal microbiome dynamics in women with recurrent bacterial vaginosis: recognition of the conversion process. *PLoS One* 8:e82599. <http://dx.doi.org/10.1371/journal.pone.0082599>.
 41. Brook I. 2002. Microbiology and management of polymicrobial female genital tract infections in adolescents. *J Pediatr Adolesc Gynecol* 15: 217–226. [http://dx.doi.org/10.1016/S1083-3188\(02\)00159-6](http://dx.doi.org/10.1016/S1083-3188(02)00159-6).
 42. Schwebke JR, Muzny CA, Josey WE. 2014. Role of *Gardnerella vaginalis* in the pathogenesis of bacterial vaginosis: a conceptual model. *J Infect Dis* 210:338–343. <http://dx.doi.org/10.1093/infdis/jiu089>.
 43. De Cáceres M, Legendre P. 2009. Associations between species and groups of sites: indices and statistical inference. *Ecology* 90:3566–3574. <http://dx.doi.org/10.1890/08-1823.1>.
 44. De Cáceres M, Legendre P, Moretti M. 2010. Improving indicator species analysis by combining groups of sites. *Oikos* 119:1674–1684. <http://dx.doi.org/10.1111/j.1600-0706.2010.18334.x>.
 45. Hill JE, Goh SH, Money DM, Doyle M, Li A, Crosby WL, Links M, Leung A, Chan D, Hemmingsen SM. 2005. Characterization of vaginal microflora of healthy, nonpregnant women by chaperonin-60 sequence-based methods. *Am J Obstet Gynecol* 193:682–692. <http://dx.doi.org/10.1016/j.ajog.2005.02.094>.
 46. Hyman RW, Fukushima M, Diamond L, Kumm J, Giudice LC, Davis RW. 2005. Microbes on the human vaginal epithelium. *Proc Natl Acad Sci U. S. A.* 102:7952–7957. <http://dx.doi.org/10.1073/pnas.0503236102>.
 47. Zhou X, Hansmann MA, Davis CC, Suzuki H, Brown CJ, Schütte U, Pierson JD, Forney LJ. 2010. The vaginal bacterial communities of Japanese women resemble those of women in other racial groups. *FEMS Immunol Med Microbiol* 58:169–181. <http://dx.doi.org/10.1111/j.1574-695X.2009.00618.x>.
 48. Hammerschlag MR, Alpert S, Rosner I, Thurston P, Semine D, McCormack WM. 1978. Microbiology of the vagina in children: normal and potentially pathogenic organisms. *Pediatrics* 62:57–62.
 49. Witkin SS, Mendes-Soares H, Linhares IM, Jayaram A, Ledger WJ, Forney LJ. 2013. Influence of vaginal bacteria and D- and L-lactic acid isomers on vaginal extracellular matrix metalloproteinase inducer: implications for protection against upper genital tract infections. *mBio* 4:e00460-13. <http://dx.doi.org/10.1128/mBio.00460-13>.
 50. Mendes-Soares H, Suzuki H, Hickey RJ, Forney LJ. 2014. Comparative functional genomics of *Lactobacillus* spp. reveals possible mechanisms for specialization of vaginal lactobacilli to their environment. *J Bacteriol* 196: 1458–1470. <http://dx.doi.org/10.1128/JB.01439-13>.
 51. Koumans EH, Sternberg M, Bruce C, McQuillan G, Kendrick J, Sutton M, Markowitz LE. 2007. The prevalence of bacterial vaginosis in the United States, 2001–2004; associations with symptoms, sexual behaviors, and reproductive health. *Sex Transm Dis* 34:864–869. <http://dx.doi.org/10.1097/OLQ.0b013e318074e565>.
 52. Workowski KA, Berman S, Centers for Disease Control and Prevention (CDC). 2010. Sexually transmitted diseases treatment guidelines, 2010. *MMWR Recomm Rep* 59:1–110.
 53. Srinivasan S, Morgan MT, Liu C, Matsen FA, Hoffman NG, Fiedler TL, Agnew KJ, Marrazzo JM, Fredricks DN. 2013. More than meets the eye: associations of vaginal bacteria with Gram stain morphotypes using molecular phylogenetic analysis. *PLoS One* 8:e78633. <http://dx.doi.org/10.1371/journal.pone.0078633>.
 54. Yeoman CJ, Thomas SM, Miller ME, Ulanov AV, Torralba M, Lucas S, Gillis M, Cregger M, Gomez A, Ho M, Leigh SR, Stumpf R, Creedon DJ, Smith MA, Weisbaum JS, Nelson KE, Wilson BA, White BA. 2013. A multi-omic systems-based approach reveals metabolic markers of bacterial vaginosis and insight into the disease. *PLoS One* 8:e56111. <http://dx.doi.org/10.1371/journal.pone.0056111>.
 55. Hickey RJ, Forney LJ. 2014. *Gardnerella vaginalis* does not always cause bacterial vaginosis. *J Infect Dis* 210:1682–1683. <http://dx.doi.org/10.1093/infdis/jiu303>.
 56. Swidsinski A, Mendling W, Loening-Baucke V, Ladhoff A, Swidsinski S, Hale LP, Lochs H. 2005. Adherent biofilms in bacterial vaginosis. *Obstet Gynecol* 106:1013–1023. <http://dx.doi.org/10.1097/01.AOG.0000183594.45524.d2>.
 57. Santiago GL, Deschaght P, El Aila N, Kiama TN, Verstraelen H, Jefferson KK, Temmerman M, Vaneechoutte M. 2011. *Gardnerella vaginalis* comprises three distinct genotypes of which only two produce sialidase. *Am J Obstet Gynecol* 204:450.e1–450.e7. <http://dx.doi.org/10.1016/j.ajog.2010.12.061>.
 58. Pybus V, Onderdonk AB. 1997. Evidence for a commensal, symbiotic relationship between *Gardnerella vaginalis* and *Prevotella bivia* involving ammonia: potential significance for bacterial vaginosis. *J Infect Dis* 175: 406–413. <http://dx.doi.org/10.1093/infdis/175.2.406>.
 59. Fredricks DN, Fiedler TL, Marrazzo JM. 2005. Molecular identification of bacteria associated with bacterial vaginosis. *N Engl J Med* 353: 1899–1911. <http://dx.doi.org/10.1056/NEJMoa043802>.
 60. Eren AM, Zozaya M, Taylor CM, Dowd SE, Martin DH, Ferris MJ. 2011. Exploring the diversity of *Gardnerella vaginalis* in the genitourinary tract microbiota of monogamous couples through subtle nucleotide variation. *PLoS One* 6:e26732. <http://dx.doi.org/10.1371/journal.pone.0026732>.
 61. Shafer MA, Sweet RL, Ohm-Smith MJ, Shalwitz J, Beck A, Schachter J. 1985. Microbiology of the lower genital tract in postmenarchal adolescent girls: differences by sexual activity, contraception, and presence of non-specific vaginitis. *J Pediatr* 107:974–981. [http://dx.doi.org/10.1016/S0022-3476\(85\)80208-0](http://dx.doi.org/10.1016/S0022-3476(85)80208-0).
 62. Myhre AK, Mykkestad K, Adams JA. 2010. Changes in genital anatomy and microbiology in girls between age 6 and age 12 years: a longitudinal study. *J Pediatr Adolesc Gynecol* 23:77–85. <http://dx.doi.org/10.1016/j.jpag.2009.05.012>.
 63. Fethers K, Twin J, Fairley CK, Fowkes FJ, Garland SM, Fehler G, Morton AM, Hocking JS, Tabrizi SN, Bradshaw CS. 2012. Bacterial vaginosis (BV) candidate bacteria: associations with BV and behavioural practices in sexually experienced and inexperienced women. *PLoS One* 7:e30633. <http://dx.doi.org/10.1371/journal.pone.0030633>.
 64. Harwich MD, Alves JM, Buck GA, Strauss JF, Patterson JL, Oki AT, Girerd PH, Jefferson KK. 2010. Drawing the line between commensal and

- pathogenic *Gardnerella vaginalis* through genome analysis and virulence studies. *BMC Genomics* 11:375. <http://dx.doi.org/10.1186/1471-2164-11-375>.
65. Yeoman CJ, Yildirim S, Thomas SM, Durkin AS, Torralba M, Sutton G, Buhay CJ, Ding Y, Dugan-Rocha SP, Muzny DM, Qin X, Gibbs RA, Leigh SR, Stumpf R, White BA, Highlander SK, Nelson KE, Wilson BA. 2010. Comparative genomics of *Gardnerella vaginalis* strains reveals substantial differences in metabolic and virulence potential. *PLoS One* 5:e12411. <http://dx.doi.org/10.1371/journal.pone.0012411>.
 66. Ahmed A, Earl J, Retchless A, Hillier SL, Rabe LK, Cherpes TL, Powell E, Janto B, Eutsey R, Hiller NL, Boissy R, Dahlgren ME, Hall BG, Costerton JW, Post JC, Hu FZ, Ehrlich GD. 2012. Comparative genomic analyses of 17 clinical isolates of *Gardnerella vaginalis* provide evidence of multiple genetically isolated clades consistent with subspeciation into genovars. *J Bacteriol* 194:3922–3937. <http://dx.doi.org/10.1128/JB.00056-12>.
 67. Paramel Jayaprakash T, Schellenberg JJ, Hill JE. 2012. Resolution and characterization of distinct cpn60-based subgroups of *Gardnerella vaginalis* in the vaginal microbiota. *PLoS One* 7:e43009. <http://dx.doi.org/10.1371/journal.pone.0043009>.
 68. Robbins CL, Fortenberry JD, Roth AM, Ott MA. 2012. Premenarchal girls' genital examination experiences. *J Adolesc Health* 51:179–183. <http://dx.doi.org/10.1016/j.jadohealth.2011.11.009>.
 69. Elsner P, Maibach HI. 1990. Microbiology of specialized skin: the vulva. *Semin Dermatol* 9:300–304.
 70. Farage M, Maibach H. 2006. Lifetime changes in the vulva and vagina. *Arch Gynecol Obstet* 273:195–202. <http://dx.doi.org/10.1007/s00404-005-0079-x>.
 71. Brown CJ, Wong M, Davis CC, Kanti A, Zhou X, Forney LJ. 2007. Preliminary characterization of the normal microbiota of the human vulva using cultivation-independent methods. *J Med Microbiol* 56: 271–276. <http://dx.doi.org/10.1099/jmm.0.46607-0>.
 72. Nugent RP, Krohn MA, Hillier SL. 1991. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of Gram stain interpretation. *J Clin Microbiol* 29:297–301.
 73. Yuan S, Cohen DB, Ravel J, Abdo Z, Forney LJ. 2012. Evaluation of methods for the extraction and purification of DNA from the human microbiome. *PLoS One* 7:e33865. <http://dx.doi.org/10.1371/journal.pone.0033865>.
 74. Hou D, Zhou X, Zhong X, Settles ML, Herring J, Wang L, Abdo Z, Forney LJ, Xu C. 2013. Microbiota of the seminal fluid from healthy and infertile men. *Fertil Steril* 100:1261–1269. <http://dx.doi.org/10.1016/j.fertnstert.2013.07.1991>.
 75. Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ. 2008. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. *Appl Environ Microbiol* 74:2461–2470. <http://dx.doi.org/10.1128/AEM.02272-07>.
 76. Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. <http://dx.doi.org/10.1186/1471-2105-9-559>.
 77. Legendre P, Legendre L. 2012. Cluster analysis, p 337–424. In *Numerical ecology*, 3rd ed. Elsevier B.V., Amsterdam, The Netherlands.
 78. Borcard D, Gillet F, Legendre P. 2011. *Numerical ecology with R*. Springer, New York, NY.
 79. Oksanen J, Blanchet FG, Kindt R, Legendre P. 2013. *Vegan: community ecology package*. R package version 2.0-10. <http://CRAN.R-project.org/package=vegan>.
 80. Rao CR. 1995. A review of canonical coordinates and an alternative to correspondence analysis using Hellinger distance. *Quaestio* 19:23–63.
 81. Legendre P, Gallagher E. 2001. Ecologically meaningful transformations for ordination of species data. *Oecologia* 129:271–280. <http://dx.doi.org/10.1007/s004420100716>.
 82. Bray JR, Curtis JT. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecol Monogr* 27:325–349. <http://dx.doi.org/10.2307/1942268>.
 83. Jain AK, Dubes RC. 1988. *Algorithms for clustering data*. Prentice Hall, Inc., Upper Saddle River, NJ, USA.
 84. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler M, Magnusson A, Moeller S. 2014. *gplots: Various R programming tools for plotting data*. R package version 2.14.1. <http://CRAN.R-project.org/package=gplots>.
 85. Cailliez F, Pages J-P, Morlat G, Amiard JC. 1976. *Introduction à l'analyse des données*. Société de mathématiques appliquées et de sciences humaines, Paris, France.
 86. Bates D, Maechler M, Bolker B, Walker S. 2014. *lme4: linear mixed-effects models using Eigen and S4*. R package version 1.1-7. <http://CRAN.R-project.org/package=lme4>.
 87. Warton DI, Hui FK. 2011. The arcsine is asinine: the analysis of proportions in ecology. *Ecology* 92:3–10. <http://dx.doi.org/10.1890/10-0340.1>.
 88. Kuznetsova A, Brockhoff PB, Christensen R. 2014. *lmerTest: Tests in linear mixed effects models*. R package version 2.0-20. <http://CRAN.R-project.org/package=lmerTest>.
 89. Dufréne M, Legendre P. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecol Monogr* 67:345–366. [http://dx.doi.org/10.1890/0012-9615\(1997\)067\[0345:SAAI\]2.0.CO;2](http://dx.doi.org/10.1890/0012-9615(1997)067[0345:SAAI]2.0.CO;2).