Short title: Using Bayesian networks to association metadata and the vaginal microbiome

Title: Associations between sexual habits, menstrual hygiene practices, demographics and the vaginal microbiome as revealed by Bayesian network analysis

Noelle Noyes[1], Kyu-Chul Cho[2], Jacques Ravel[3], Larry J. Forney[4], Zaid Abdo[1*]

[1] Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, Colorado 80523, USA

[2] Department of Statistics, University of Idaho, Moscow, Idaho 83844. Current Address: Naver Corporation, Jeong-ja-dong 178-1, Bundang-gu, Seongnam-si, Gyeonggi-do, South Korea

[3] Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore Maryland 21201, USA

[4] Department of Biological Sciences, University of Idaho, Moscow, Idaho 83844, USA

[*] Corresponding author

E-mail: zaid.abdo@colostate.edu

Abstract

The vaginal microbiome plays an influential role in several disease states in reproductive age women, including bacterial vaginosis (BV). While demographic characteristics are associated with differences in vaginal microbiome community structure, little is known about the influence of sexual and hygiene habits. Furthermore, associations between the vaginal microbiome and risk symptoms of bacterial vaginosis have not been fully elucidated. Using Bayesian network (BN) analysis of 16S rRNA gene sequence results, demographic and extensive questionnaire data, we describe both novel and previously documented associations between habits of women and their vaginal microbiome. The BN analysis approach shows promise in uncovering complex associations between disparate data types. Our findings based on this approach support published associations between specific microbiome members (e.g., *Eggerthella*, *Gardnerella*, *Dialister*,

Sneathia and *Ruminococcaceae*), the Nugent score (a BV diagnostic) and vaginal pH (a risk symptom of BV). Additionally, we found that several microbiome members were directly connected to other risk symptoms of BV (such as vaginal discharge, odor, itch, irritation, and yeast infection) including *L. jensenii*, *Corynebacteria*, and *Proteobacteria*. No direct connections were found between the Nugent Score and risk symptoms of BV other than pH, indicating that the Nugent Score may not be the most useful criteria for assessment of clinical BV. We also found that demographics (i.e., age, ethnicity, previous pregnancy) were associated with the presence/absence of specific vaginal microbes. The resulting BN revealed several as-yet undocumented associations between birth control usage, menstrual hygiene practices and specific microbiome members. Many of these complex relationships were not identified using common analytical methods, i.e., ordination and PERMANOVA. While these associations require confirmatory follow-up study, our findings strongly suggest that future studies of the vaginal microbiome and vaginal pathologies should include detailed surveys of participants' sanitary, sexual and birth control habits, as these can act as confounders in the relationship between the microbiome and disease. Although the BN approach is powerful in revealing complex associations within multidimensional datasets, the need in some cases to discretize the data for use in BN analysis can result in loss of information. Future research is required to alleviate such limitations in constructing BN networks. Large sample sizes are also required in order to allow for the incorporation of a large number of variables (nodes) into the BN, particularly when studying associations between metadata and the microbiome. We believe that this approach is of great value, complementing other methods, to further our understanding of complex associations characteristic of microbiome research.

Introduction

The microbiome plays a critical role in human health, and the vaginal microbiome has been linked to urogential diseases of reproductive age women, including bacterial vaginosis (BV) (1–3). BV affects nearly one-third of women in the United States (4) and has been implicated in poorer pregnancy outcomes, acquisition of sexually-transmitted infections and other vaginal disorders (5,6). Specific changes in the vaginal microflora are associated with BV, including a depletion of *Lactobacillus* species and an increased abundance of strictly anaerobic bacteria (7). However, no single bacterial taxon has been shown to cause BV and the condition can be found in women with widely varying vaginal microbiomes (8–11). BV is characterized clinically by itching, pain, burning, odor and/or discharge, and is often diagnosed based on a combination of symptoms, vaginal pH and cytological findings (12). BV is also diagnosed using the Nugent score, which is the most commonly used diagnostic test for BV within the research community (13).

Just as the microbial underpinnings of BV are complex and varied (14), so too are the influences of a woman's sexual, sanitary and other practices. A wide range of factors have been shown to increase BV risk, including smoking, douching, menstruation, and new sexual partners (15–18). Different ethnic groups exhibit varying rates of BV, leading to conclusions that intrinsic host factors may contribute to the condition (4), although the effect of confounders on this association has been questioned (19). Further complicating this picture is evidence that women from different ethnic groups tend to harbor different vaginal microflora, independent of BV status (8,20).

77 BV is therefore a multifactorial disease mediated by a complex interplay of host, microbial and
78 environmental factors. Fortunately, all three of these influences can be measured and evaluated
79 using a combination of 16S rRNA gene sequencing of vaginal samples and detailed
80 questionnaires of study participants. However, the analytical methods for identifying associations
81 between multivariate, often categorical metadata with counts of microbial taxa are either ill-
82 suited, contested or opaque (21,22). The most widely used analytical method for community-
83 level data involves ordination of normalized microbiome counts (e.g., principal coordinates
84 analysis [PCoA] and nonmetric multidimensional scaling [NMDS]) followed by statistical
85 significance testing of explanatory/metadata variables (e.g., ANOSIM, PERMANOVA).
86 However, these methods reduce the dimensionality of the microbial community structure to two
87 or three dimensions, obscuring intra-microbiome interactions. In addition, interactions between
88 explanatory (metadata) variables are difficult or impossible to uncover because statistical testing
89 usually occurs on a variable-by-variable basis (23). Therefore, these methods do not allow for
90 discovery of more nuanced and complex dynamics between and within the microbiome and host
91 and environmental factors. Identifying such dynamics can be crucial for understanding and thus
92 combatting multifactorial conditions such as bacterial vaginosis – particularly when such
93 conditions lack an accurate diagnostic description and/or test (24). Bayesian network (BN)
94 analysis offers potential advantages in handling mixed datasets for complex conditions, and use
95 of this approach may help to elucidate some of the nuanced, yet important, associations between
96 host, environmental and microbial factors in diseases such as BV. For instance, BN's have been
97 proposed as diagnostic tools for such multifactorial diseases as breast cancer and cardioembolic
98 stroke (25–27). While the heavy computational burden of building BNs has impeded their
99 widespread adoption, recent advances in algorithms have removed this barrier (28). In addition,

software implementations now offer BN analysis within user-friendly packages (29,30). Finally, unlike other machine learning approaches, BN's possess an inherently intuitive interpretation and allow for a wide range of data input types, both continuous and categorical (31).

The primary goal of this study was to uncover associations between host behavioral characteristics, vaginal microbiome composition, risk symptoms and diagnostic criteria of BV utilizing BN's. Using this approach, we have demonstrated associations between women's sexual and menstrual habits, demographics, vaginal microbiome composition, risk symptoms of BV and the Nugent Score (a BV diagnostic). Our findings support previously-documented associations between microbiome members (e.g., *Eggerthella*, *Gardnerella*, *Dialister*, *Sneathia* and *Ruminococcaceae*), the Nugent Score and vaginal pH (risk symptom) (8,13,32,33). However, we found no connections between the Nugent Score and other risk symptoms of BV such as vaginal discharge, itch, irritation, abdominal or pelvic pain, yeast infection and underwear staining within 60 days prior to sampling and any type of current vaginal odor. This suggests that the Nugent Score may not provide an accurate diagnostic of BV in some women. Additionally, we found that several microbiome members were directly connected to risk symptoms of BV, including *L. jensenii*, *Corynebacteria*, and *Proteobacteria*. Demographics (i.e., age, ethnicity, previous pregnancy) also influenced the presence/absence of specific vaginal microbes. The resulting network also revealed several as-yet undocumented associations between birth control usage, menstrual hygiene practices and microbiome members. These associations require confirmatory follow-up study, though strongly suggest that future studies of the vaginal microbiome and vaginal pathologies include detailed surveys of participants' sanitary, sexual and

122 birth control habits, as these can act as confounders in the relationship between the microbiome
123 and disease.
124
125 Secondary objectives of this study included demonstrating the hypothesis-generating power of
126 BN's through confirmation of previously described microbiome-BV associations, as well as the
127 wide accessibility of BN's within a readily available, user-friendly environment. In addition, we
128 aimed to illustrate the capacity of the BN method to directly associate component taxa of the
129 microbiome with metadata. This capacity contrasts with common approaches used to analyze
130 these data types (e.g., PCoA, NMDS and clustering), which rely on reducing the dimensionality
131 of the total microbial community to two or three dimensions, followed by clustering or
132 correlation of this reduced community with metadata (such as demographics). The BN approach
133 also circumvents the limits on the number of interactions that can be tested between metadata
134 and microbiome members that are imposed by tools such as PERMANOVA (34) or
135 metagenomeSeq (35). This advantage of the BN approach is especially true when the number of
136 variables observed in association with the microbiome is large. In such cases, network analysis
137 identifies the hierarchy of relationships between the various metadata and microbiome taxa (i.e.,
138 the nodes of the network), and then represents these relationships using the arcs (or edges) of the
139 network. When possible, we demonstrated these differences by providing direct comparisons of
140 the BN approach with results obtained from Nonmetric Multidimensional Scaling Ordination
141 (NMDS)-Analysis of Similarity (ANOSIM) and PERMANOVA. These comparisons should be
142 interpreted with caution, as particularly the NMDS/ANOSIM approach is limited by the degrees
143 of freedom available for significance testing and for assessing interactions. In addition -- and

unlike the BN approach -- NMDS/ANOSIM and PERMANOVA analyses do not allow for the possibility of hierarchical relationships between variables observed within the data.

It is important to note that the BN approach presented in this work was aimed at inference, rather than prediction; and that we focused on evaluating associations within the data that could indicate possible associations in the target population. Use of the BN approach in this way does not attempt to classify the state of a new individual within the target population, i.e., prediction. Hence our network was fitted using all data and we did not attempt to perform cross validation to assess the ability of the network to predict or classify.

Materials and Methods

Study Population and Sampling

Questionnaire responses (see S1 File) and the relative abundance of 16S rRNA gene sequences (see S1 File) were obtained from a study of 396 healthy, sexually active, non-pregnant women of reproductive age (range 12-45 years); details of the study population have been described previously (8). Briefly, study participants submitted 2 self-collected vaginal swabs; one swab was used for 16S sequencing and the other was used to obtain a Nugent score, which is one of several diagnostic assays for BV and is the most commonly used within the research community (13). High Nugent scores are considered to be diagnostic of BV. Since all study participants were considered to be healthy, those with high Nugent scores were assumed to have asymptomatic BV. At the same time as vaginal swab collection, study participants were

164 administered a detailed questionnaire on their sexual and sanitary habits and health histories (8,
165 S2 File).

166 Ethics Statement

167 The institutional review boards at Emory University School of Medicine, Grady Memorial
168 Hospital and the University of Maryland School of Medicine approved the protocol. Guidelines
169 of the universities were followed in the conduct of the clinical research. Participants provided
170 written informed consent to participate in the study and written informed consent for use of the
171 data for future studies. The study was registered at clinicaltrials.gov under ID NCT00576797.

172 Data Preparation

173 Several filtering steps were taken with the goal of retaining only those variables (i.e., taxa,
174 survey responses and demographic information) that would provide robust information during
175 BN construction. In order to reach this goal, we removed taxa and survey response variables that
176 were sparsely represented within the dataset, as described below.

177 Ribosomal 16S sequencing data were processed using the Ribosomal Database Project (RDP)
178 Classifier (36) as described previously (8). In order to allow for robust BN construction, we
179 removed samples that contained <1,000 16S counts. In addition, subjects who were not in overall
180 general health or who had experienced toxic shock syndrome were removed from analysis. 16S
181 counts were then normalized to the sample with the lowest number of counts, and taxa with
182 counts <0.1% of total 16S counts were removed from further analysis to provide robust inputs
183 into BN analysis. Finally, 16S counts were converted into presence/absence (i.e., 0/1) data (S3

File). While the implementation of BN analysis used in this study, i.e., bnlearn (29), is able to incorporate continuous variables with a Gaussian distribution, unfortunately the sparseness and skewness of the 16S rRNA count distributions in this dataset were such that we could not justify using abundances without dichotimization (see abundance distributions of used taxa in S8 File). Furthermore, categorizing these distributions to include more than presence/absence categories would have inflated the number of parameters that needed to be estimated, resulting in diminished stability of the BN network given the available sample size.

From the survey data, we removed questions that were related solely to study design factors (e.g., enrollment information, location of survey administration, survey administrator ID). Binary questions about specific vaginal odors (i.e., fishy, musty, foul and other) were collapsed into a single, binary "any odor" variable in order to decrease sparseness. In addition, we dropped questions about past birth control use in order to focus on current birth control habits, which we hypothesized would exhibit a stronger effect on the vaginal microbiome. Survey questions for which >5% of respondents did not provide an answer were excluded from BN analysis so that variables with a significant proportion of missing data did not exert undue influence over the graph. For the same reason, and after this variable-level exclusion, we removed any participants with missing data, keeping only those subjects with complete data in conjunction with the variables under study. After these filtering steps, we discretized any remaining continuous variables. Respondents' ages were converted from a continuous into a categorical variable with 3 levels: less than or equal to 30 (young adults), 31 – 40, and greater than 40 years of age (close to menopause). Nugent scores were categorized into low (0-3), medium (4-6) and high (7-10). Age at menstrual onset was discretized into <11, 11-15 and >15 years, and number of sexual partners in the last 60 days was categorized into 0, 1 and >1. Number of vaginal births and duration of

menstruation were treated as categorical variables with 4 levels each (0 – 3 births and 1 – 4 days, respectively). Finally, we excluded survey questions with sparse outcomes, defined as having <5% of respondents within any one outcome level (i.e., for binary questions, at least 5% of respondents had to be in the "yes" and "no" outcome levels). This was also done such that stable BN's could be constructed. The resulting data are presented in S3 File and a key is available from S4 File.

BN Construction

Filtered, dichotomized 16S rRNA gene counts as well as filtered and discretized demographic and survey responses were provided as input to BN construction, which was performed using bnlearn (29) implemented within the statistical software R (37). Age and ethnicity were specified as roots of the network, and Nugent score was specified as a leaf. Directionality between all other nodes was not specified in order to allow for the complex, likely multidirectional interplay between the microbiome, host factors, vaginal microenvironment and risk symptoms that might be associated with BV (12,14,38). The hill-climbing algorithm was used for BN construction with scoring based on the logarithm of the Bayesian Dirichlet equivalence (BDe) score; an optimal imaginary sample size was estimated using the alpha.star function in bnlearn (29,39). To obtain a consensus network, 1,000,000 BN's were constructed by first generating 1,000 random number seeds and, then, generating 1000 bootstraps of the input data. The aim of varying the seeds was to alleviate the possibility of being stuck in a local optimum in the space of all possible networks. The 1000 bootstraps per different starting point were used to assess stability of the network given each starting point. All 1,000,000 networks were then combined to construct a consensus BN by first using the mean function to pool the

229 bootstrapped networks and then by using the averaged.network function to construct the
230 consensus. Both of these functions are part of the bnlearn package. The empirical frequency for
231 every arc within all BN's was determined, and arcs that met an empirical threshold frequency of
232 0.30 or more were used to produce the consensus BN. Network analysis and inference was
233 performed using both this threshold and a 0.50 majority rule threshold. The resulting DAG was
234 used in maximum likelihood estimation parameter fitting through the bn.fit function, as well as
235 network analysis (29).

236 BN Analysis

237 Prior to BN analysis, nodes were categorized as either "demographic", "microbiome",
238 "sexual/sanitary habit", "BV risk symptom", or "BV diagnostic criteria" (S4 File). Variables
239 that were categorized as BV risk symptoms included "staining of underwear", "vaginal
240 discharge", "vaginal odor", "vaginal itch", "vaginal irritation", "abdominal or pelvic pain" and
241 "yeast infection" within the past 60 days prior to sampling, and pH and any type of current
242 vaginal odor (any odor). These variables where chosen based on previously observed
243 associations with increased risk of BV (6,12,40). Vaginal pH, vaginal discharge and staining,
244 and odor could be thought of as proxies for the Amsel symptoms used in BV diagnosis in clinical
245 settings (32). Given that all included subjects were considered healthy at the time of this study,
246 our intention was to highlight possible increased risk of BV due to presence of these risk
247 symptoms. The Nugent Score was the only variable designated as a "BV diagnostic criteria".
248 After *a priori* categorization, the nodes and edges obtained from bnlearn were used to conduct
249 network analysis in order to understand the structure of the overall network, as well as the
250 connections within it. Average Markov blanket size and overall graph characteristics were

251 calculated and reported. Network density was defined as the proportion of all possible edges that
252 occurred in the final network. Node degree was defined as the number of incoming and outgoing
253 edges, node closeness centrality was defined as the average shortest path from a node to all other
254 nodes, and node betweenness centrality was defined as the frequency with which a node was
255 included in the shortest path between nodes in the network. Node closeness and betweenness
256 centralities were normalized for comparison purposes. Groupings of nodes (i.e., groupings not
257 related to the *a priori* node category, but rather agnostic groupings) were detected using a
258 modularity optimizing algorithm (41) as implemented in Gephi (42), using a resolution of 1.0
259 and randomization without edge weights. Networks were visualized using Gephi (42), using the
260 Force Atlas layout unless otherwise specified.

261 Ordination and Clustering

262 Filtered, normalized, non-dichotomized, Hellinger-transformed (43) 16S rRNA gene counts were
263 used to construct a Euclidean distance matrix, which was then ordinated using NMDS with
264 multiple random starts and two dimensions (44). Use of two dimensions resulted in a stress value
265 of 0.19, indicating less-than-ideal fit; however, increasing the number of dimensions resulted in
266 failure of convergence, and therefore a better fit could not be achieved. Associations between the
267 ordinated microbiome and metadata variables (i.e., host demographics and survey data) were
268 tested using analysis of similarities (ANOSIM) (45). A *P*-value of less than 0.05 was considered
269 statistically significant.

270 PERMANOVA

Permutation based multivariate analysis of variance (PERMANOVA) was also used to assess the influence of metadata on the vaginal microbiome. The process of assessing significance of association followed a forward elimination, stepwise model selection approach based on Akaike's Information Criteria (AIC) and using the step function as implemented in R. PERMANOVA was performed onHellinger-transformed microbiome counts. All survey and demographic response variables that passed the stepwise selection criteria (see above) and minimized the AIC score were included in the final PERMANOVA model (implemented using rda function in vegan). A permutation test using 1000 permutations was used to compute *P*-values to assess significance of associations between the metadata and the observed microbiome.

Results and Discussion

Data filtering and descriptive statistics

Of the 396 participants included in the original study, 2 were removed *a priori* due to a lack of any metadata from the questionnaire. Among the remaining 394 participants, 247 genus- and species-level taxa were identified through 16S sequencing of vaginal swabs. Total 16S counts per sample ranged from 693 to 7,392 with a median of 2,149. Five samples with <1,000 total 16S counts were removed from the analysis, for a remaining 389 samples; all 247 taxa were represented within these 389 samples. 16S counts were normalized to a count of 1,006, which was the total number of 16S reads in the sample with the lowest total count. The distribution of 16S counts was strongly right-skewed, and therefore 220 taxa were present at <0.1% abundance

291 (i.e., fewer than an average of 2 reads per sample) and were removed from further analysis.
292 From the 389 study participants with 16S data that passed filtering, a further 4 participants were
293 removed because they answered "no" or did not answer the question of whether they were in
294 overall good health; and a further 9 were removed because they either answered "yes" to
295 previous toxic shock syndrome or did not answer this question. A further 91 respondents were
296 excluded from analysis due to missing data for survey questions. The decision to remove
297 subjects with missing data was not made lightly and was based on two reasons: 1) data
298 imputation, although useful, could be faulty or could reduce the variability of the estimates of the
299 network parameters (46) resulting in overconfidence in the network and 2) adding the missing
300 data as a category by itself would substantially hamper interpretability of results. Accordingly,
301 the decision was made to err on the conservative side (i.e., increased variability), rather than on
302 the side of overconfidence or weak interpretability. This left a total of 285 healthy study
303 participants for inclusion in BN analysis, encompassing a total of 27 taxa with a median
304 prevalence of 89/285 samples (31%, range 24 - 228 out of 285 samples, S3 File). Among the
305 285 participants included in the network, 58 self-reported to be Hispanic (20.4%), 71 Asian
306 (24.9%), 78 White (27.4%) and 78 African-American (27.4%). Nearly half of respondents were
307 age 30 or under (49.5%, 141/285), and only 25 reported being over age 40 (8.8%).

308 After removing study-specific questions, redundant questions, past birth control use questions,
309 questions about frequency of sexual habits and menstrual hygiene, as well as questions related to
310 study exclusion criteria, a total of 263 survey questions were included in the initial survey
311 response data. After removing questions with >5% missing responses, we were left with 56
312 study questions. A further 27 questions were removed due to sparseness, resulting in a total of 29
313 study questions available for input into the BN (S3 File). A majority of the 285 survey

respondents included in the Bayesian network reported having been pregnant at least once (170/285, 59.6%). Sixty-three reported having had no sex partners in the 60 days preceding sampling (22.1%), a further 199 were monogamous in this same time period (69.8%), and 23 respondents had multiple sex partners (8.1%). The majority of women reported an absence of vaginal itching, irritation, discharge, odor or uro-abdominal pain in the 60 days preceding the survey (177/285, 62.1%). Vaginal yeast infections in the 60 days preceding sampling were reported by 19 of the 285 respondents. In addition to the 29 survey questions, we also had data on the vaginal pH and Nugent score of the 285 women in the study. The majority of participants had a low Nugent score (188/285, 66.0%), while a smaller proportion were categorized as intermediate and high (34/285 and 63/285, 11.9% and 22.1%, respectively). Around half of the 285 women had a vaginal pH of less than or equal to 4.5 (152/285, 53.3%), while about half had a higher vaginal pH (133/285, 46.7%).

Network construction

A total of 60 variables were included in BN construction: the 29 survey response variables and 27 bacterial taxa that passed filtering criteria, as well as age, ethnicity, vaginal pH and Nugent score. The consensus BN had 152 directed edges with an average Markov blanket size of 9.4.

BN Characteristics

The final consensus graph contained 60 nodes within a single sparse network with a density of 0.043. The majority of nodes contained between 1 and 6 connections, while several were more highly connected (Fig. 1). *Eggerthella* represented the most highly-connected node with 11 edges, followed by *Sneathia*, previous pregnancy status and current use of any type of birth control with 10 edges each.

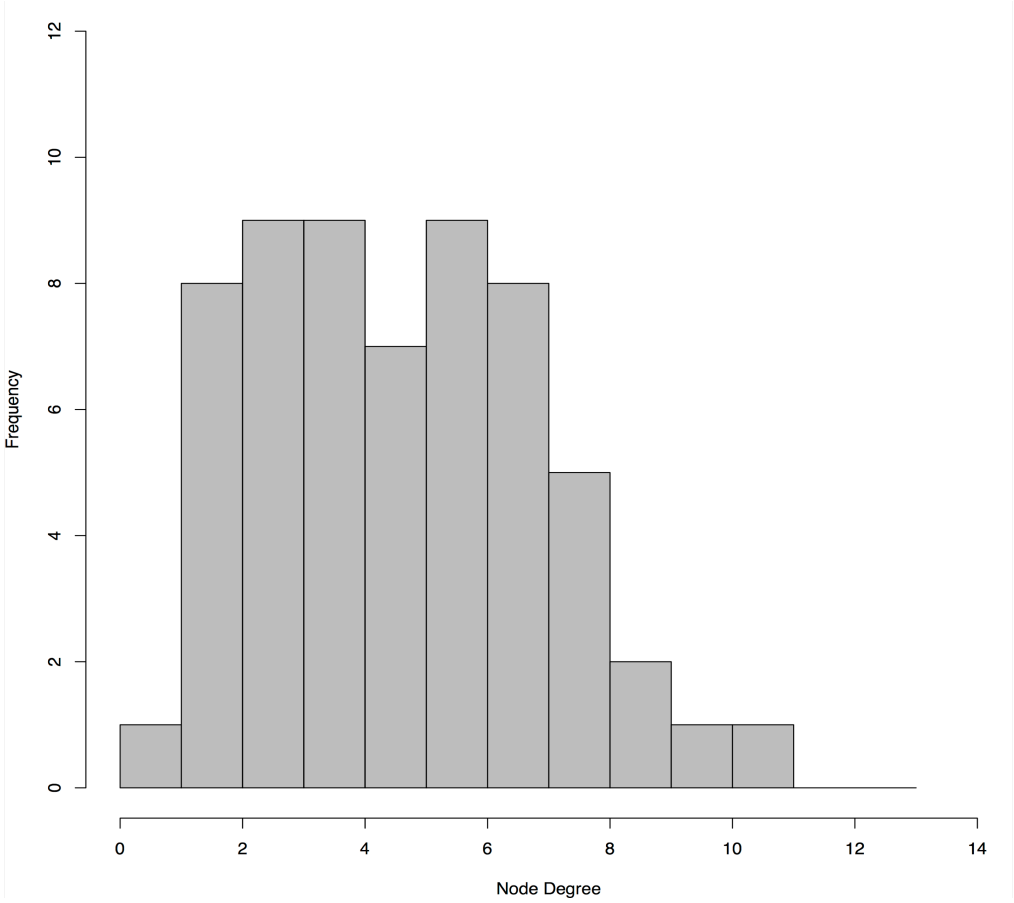


Figure 1. Histogram of node degree for the final BN. The majority of nodes had fewer than 8 connections (i.e., degrees), while a few were more highly connected.

Betweenness centrality measures indicated that *Parvimonas*, report of vaginal odor in the 60 days prior to the survey, as well as presence/absence of *L. vaginalis* within the vaginal microbiome, were situated along highly-connected paths in the network [S4 File, see column

341 "Betweenness Centrality"]. Node closeness centrality metrics suggested that presence/absence
342 of *Lactobacillus iners, Mycoplasmataceae, Lachnospiraceae* and *Streptococcus* were all nodes at
343 the geographic center of the network, along with vaginal irritation in the 60 days preceding the
344 survey [S4 File, see column "Closeness Centrality"]. Interestingly, *Sneathia*, *Eggerthella*,
345 *Parvimonas* and *Streptococcus* have been described as key members of a vaginal microbiome
346 characterized by a diversity of non-*Lactobacillus* bacteria found more commonly in black and
347 Hispanic women (8,47). The fact that these bacteria are centrally situated in the BN developed in
348 this study (as defined by closeness, betweenness and degree metrics) lends support to the idea
349 that these bacteria could play an important role in differentiating the vaginal microbiome in
350 different women.

351 Ordination, Clustering and PERMANOVA

352 Univariable clustering analysis with ANOSIM significance testing revealed that six of the 33
353 metadata variables (including host demographics and survey responses) were statistically
354 significantly associated with the microbiome ordination results. This list comprised ethnicity,
355 vaginal pH, Nugent score, tampon use, as well as previous pregnancy and vaginal birth statuses
356 [S5 File]. S6 File provides the associated NMDS plots for these six variables. These analyses
357 suggested that demographic and behavioral traits of participants were associated with differences
358 in the microbiome; however, given the nature of these analyses, it was unclear specifically which
359 microbiome members were being affected, and whether such associations were direct or indirect.
360 PERMANOVA analysis, alternatively, revealed a best-fit model that also included vaginal pH,
361 Nugent score and previous pregnancy, as well as participant age, use of pads during
362 menstruation, staining of underwear in the previous 60 days, and use of male condoms as current

363 birth control method, which together explained 25% of the variability in the data (Table 1). This
364 type of analysis provided more information highlighting associations between high Nugent
365 Score, *Sneathia* and *Megasphaera* and possibly *Atopobium*, *Dialister*, *Eggerthella*,
366 *Ruminococcaceae*, and *Lachnospiraceae* (S7 Figure). It also showed putative association
367 between intermediate Nugent Score and *L. gasseri* (S7 Figure). Other relationships were tenuous
368 and not clearly defined. Furthermore, the above analyses precluded the testing of more
369 complicated interactions between multiple variables of diverse types. BN analysis was thus used
370 to identify such relationships within all of the data, including between individual microbiome
371 members; we point to the results of NMDS/ANOSIM and PERMANOVA where relevant.

372 Associations between population demographics, the microbiome and 373 survey responses

374 There were 58 edges that bridged nodes of different types, i.e., demographics, sexual and
375 menstrual hygiene habits, risk symptoms of BV, diagnostic criteria and microbiome (see S4 File
376 for *a priori* classification of nodes); compared to 94 edges that connected nodes of like types. Of
377 these 58 bridging edges, 16 were identified in 50% of the bootstrapped BN networks (Fig. 2).
378 Many of these connections confirmed previously documented relationships between
379 demographic factors and sexual and menstrual habits. For instance, ethnicity was directly related
380 to Pap testing status, with Asian women exhibiting a much higher likelihood of never having had
381 a pap smear compared to Hispanic, White and Black respondents (20.8% versus 4.6%, 2.9%, and
382 1.2%, respectively), a disparity that has been described previously (48,49). At a threshold of at
383 least 50% bootstrap support, ethnicity also influenced tampon and pad use during menstruation

(50), while at 30% support ethnicity was associated with age at menstrual onset (51,52) and the likelihood of a woman ever having self-treated a vaginal infection.

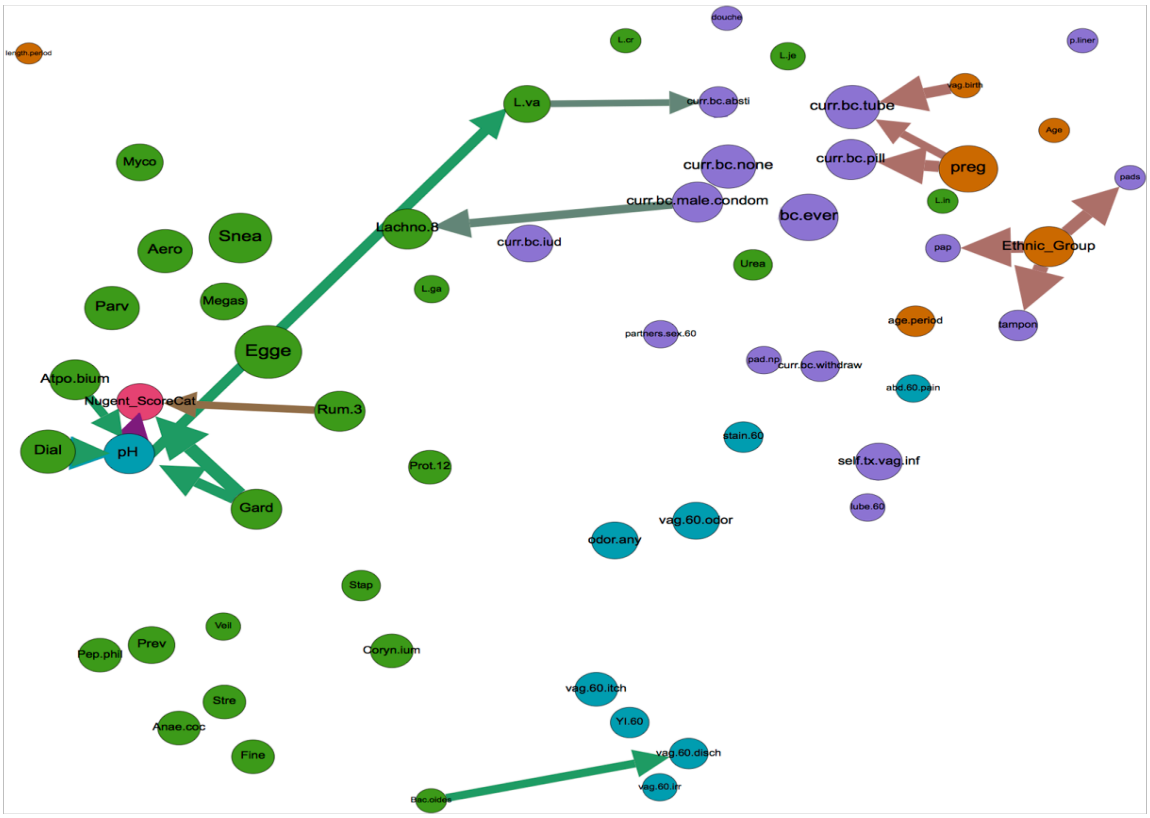
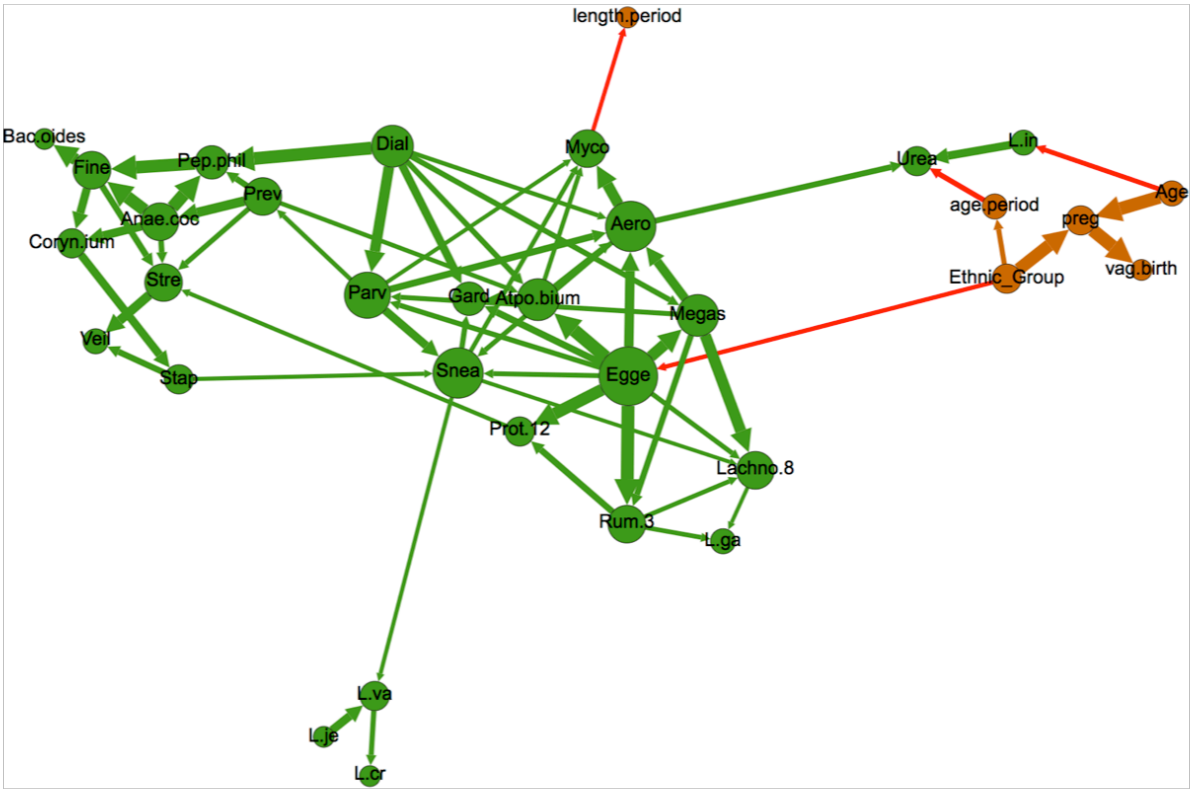


Figure 2. Inter-category bridging edges with >50% bootstrap. Network displaying all nodes, but depicting only inter-category bridging edges with >50% bootstrap support (arrows). Node size is proportional to closeness centrality, arrow thickness is proportional to bootstrap support, and node color signifies category type (orange = demographic, green = microbiome, turquoise = BV risk symptom, purple = sexual/menstrual habits, and pink = diagnostic criteria of BV). Node label abbreviations and categorization by type can be found in S4 File.

393 At 50% bootstrap support, there were no edges between demographic variables (i.e., age,
394 ethnicity, age at menstrual onset, previous pregnancy and vaginal birth) and microbiome nodes
395 (Fig. 2), suggesting that these factors may not exert robust influence over the presence of
396 common microbiome members. However, decreasing bootstrap support to 30% revealed four
397 such edges, namely associations between participant age and *Lactobacillus iners*, ethnicity and
398 *Eggerthella*, age at menstrual onset and *Ureaplasma*, and *Mycoplasmataceae* and period length
399 (Fig. 3). PERMANOVA testing supported the associations between the vaginal microbiome and
400 age, while NMDS/ANOSIM analysis supported the association between the microbiome and
401 ethnicity and marginally (ANOSIM $P = 0.06$) between the microbiome and age at menarch
402 (Table 1, S5-S7 Files). Additionally, PERMANOVA and NMDS/ANOSIM analyses revealed
403 associations between the vaginal microbiome and previous pregnancy and/or previous live birth
404 (Table 1, S5-S7 Files). Given the constraints of PERMANOVA and NMDS/ANOSIM analysis,
405 we were unable to identify which specific microbiome members may be involved in these
406 associations.



407

408 **Figure 3. Network depicting nodes related to demographics (orange) and microbial taxa**
409 **(green), as well as all edges with at least 30% bootstrap support.** Node size is proportional to
410 node degree (i.e., number of incoming and outgoing edges). Arrow thickness is proportional to
411 bootstrap support. Red arrows bridge demographic and microbiome nodes, while green and
412 orange arrows connect nodes of the same type (i.e., microbiome-to-microbiome or demographic-
413 to-demographic). Node label abbreviations can be found in S4 File.

414 However, BN analysis revealed that the probability of harboring *L. iners* given age >40 years
415 was 93%, compared to 72% and 85% for women ages 30-40 and <30 years, respectively. To
416 date, longitudinal studies of the vaginal microbiome have been restricted to relatively short time
417 periods, with subsequently little knowledge about how the microbiome may change with age.
418 However, studies suggest that variations in estrogen levels influence presence/abundance of

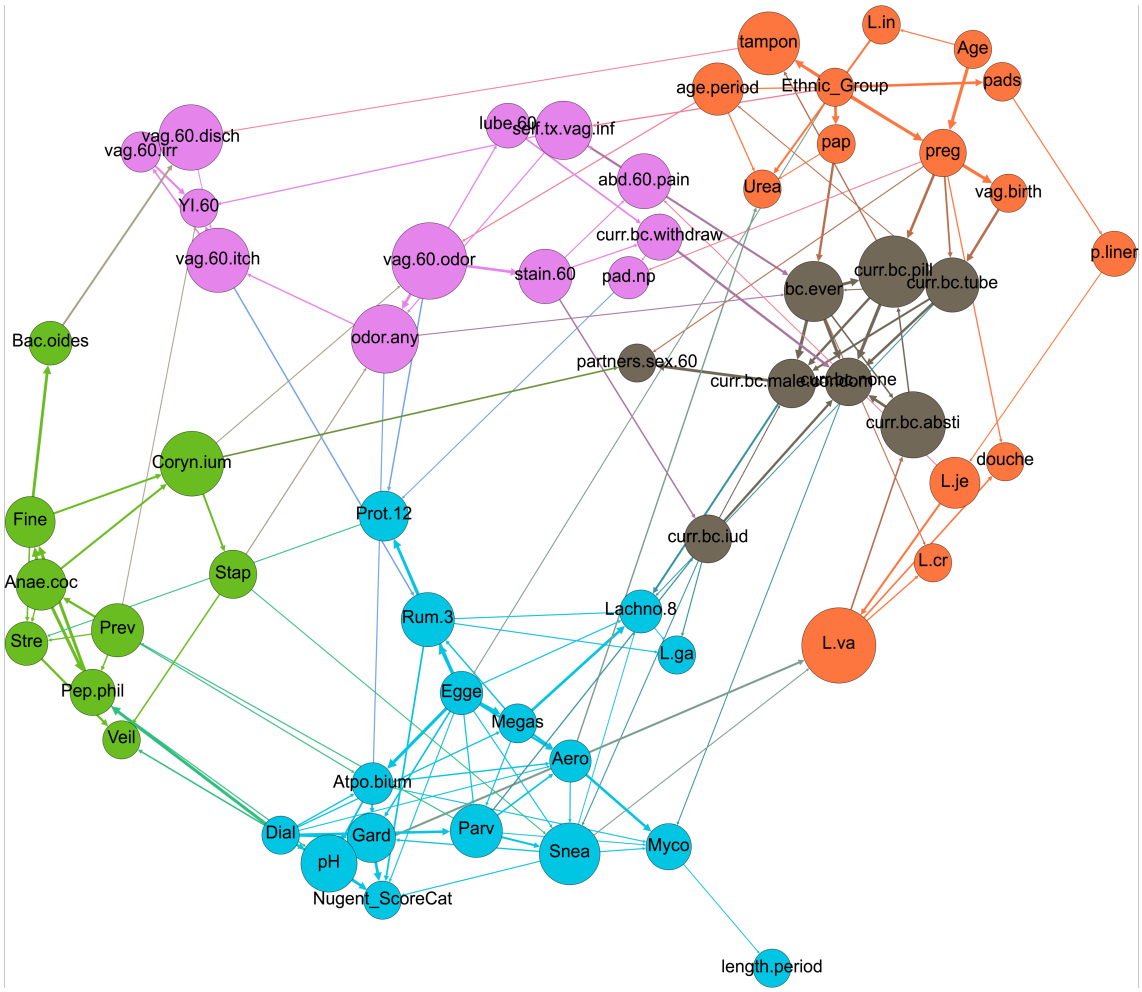
419 *Lactobacilli* (53,54), and age has a significant influence on estrogen levels (55), which together
420 present a possible mechanism for the association we found between age and *L. iners*. Existence
421 of an edge between these two nodes suggests significance of association (affirmed by a Chi-
422 square test-of-independence, p-value 0.0067). BN analysis also revealed a complex relationship
423 between age and other demographic variables. Age was associated with previous pregnancy
424 status (which can also affect estrogen levels), as was ethnicity; indeed, previous pregnancy was a
425 common effect of age and ethnicity within this dataset (i.e., the three nodes formed a v-structure
426 within the graph). Only 24% of white respondents reported having ever been pregnant (19/78),
427 compared to 58%, 77% and 86% of Asians, African-Americans and Hispanics, respectively. The
428 role of age in likelihood of previous pregnancy was very strong among Asian and white
429 respondents, and less so for African-Americans and Hispanics. Ethnicity also influenced the
430 likelihood of identifying *Eggerthella* within the microbiome, as African-Americans were more
431 likely to harbor this bacterium (45.4% versus 18.8%, 16.5% and 3.8% for Hispanics, Caucasians
432 and Asians, respectively). Interestingly, *Eggerthella* was one of 28 taxa that were previously
433 found to exhibit a significant interaction with race and BV status (47), and was also a key
434 member of a certain vaginal microbiome composition that was more commonly found in black
435 and Hispanic women than white and Asian women (8). Finally, previous pregnancy decreased
436 the likelihood of a woman harboring *L. crispatus* and age at menarch was associated with
437 *Ureasplasma* presence. *Ureaplasma* presence was also influenced by pap smear status.
438 *Ureaplasma* in the upper genital tract has been associated with poor pregnancy outcomes, and
439 *Ureaplasma* colonization in the lower genital tract has been found to be associated with a variety
440 of socioeconomic and demographic factors including educational, income, ethnicity and marital
441 status (56–58). Interestingly, age at menarch has also been associated with socioeconomic-

442 related factors such as body mass index (BMI) and ethnicity (59,60), as has pap smear status
443 (48,49). Results such as these demonstrate the complexity of interactions between microbiome
444 composition and demographic and behavioral factors, and strongly suggest that any attempts to
445 associate microbial composition with clinical BV should include co-analysis of potential
446 confounders. The results presented here indicate that age, ethnicity and previous pregnancy
447 status could potentially be used as portmanteau variables for such confounders.

448 Upon modularity analysis, nodes within the network tended to segregate into groupings of nodes
449 of like type (called "communities" in network analysis), with some exceptions (Fig. 4). Nearly
450 all of the microbiome nodes fell into two largely bacteria-specific groupings, one of which
451 contained the nodes for Nugent score and vaginal pH. All of the *Lactobacilli* except *L. gasseri*
452 segregated along with nodes related to demographics (including study participant age, ethnicity,
453 pap smear status, age at menstrual onset and previous pregnancy), menstrual hygiene habits and
454 douching. All of the nodes related to risk symptoms associated with increased risk of BV, except
455 for pH, fell into one grouping that also included use of the withdrawal method for birth control,
456 and all of the other birth control variables were clustered into a fifth grouping that also included
457 the number of sexual partners in the previous 60 days (Fig. 4).

458 These modularity-based grouping results highlight the dissociation of the Nugent score (a BV
459 diagnostic) and vaginal pH from all other risk symptoms of BV, a finding which supports recent
460 studies suggesting that these may not be the most useful criteria for assessment of clinical BV
461 disease (8,33). None of the other risk symptom-based nodes fell into the same modularity
462 grouping as the Nugent score and vaginal pH, and indeed the BN contained no directed paths of
463 any length between vaginal pH or Nugent score and risk symptoms of vaginal irritation, itch,

464 odor, pain or staining. Furthermore, contingency table propagation showed that women
465 reporting vaginal odor, itch and irritation had an 11.6% chance of a high Nugent score, compared
466 to 11.0% for women who did not report vaginal odor, itch and irritation, again suggesting that
467 the Nugent score might not be reliably associated with these risk symptoms. In addition, while
468 bacteria traditionally thought to be determinants of the Nugent score did belong to the same
469 modularity grouping as the node for Nugent score (e.g., *Dialister* and *Gardnerella*), these
470 bacteria were not associated with BV risk symptoms other than pH (Fig. 4).



471 **Figure 4: Network depicting modularity-based communities (node colors) as determined by**
472 **a modularity optimization algorithm (41)**. Size of node is proportional to node's betweenness
473 centrality. Node label abbreviations can be found in S4 File.

474 Associations between pH, Nugent Score and the microbiome

475 Several microbial taxa nodes were directly associated with vaginal pH and the Nugent Score.
476 Vaginal pH is one of the four Amsel criteria used to diagnose BV (32). A vaginal pH of less
477 than or equal to 4.5 is considered normal, while >4.5 is considered to contribute to BV.
478 Although all women in this dataset were healthy and did not have clinical BV as determined by
479 the Amsel criteria, nearly half had an "abnormal" vaginal pH >4.5 (133/285, 46.7%), a finding
480 which calls into question the use of vaginal pH as an indicator of clinical BV. As with previous
481 studies (61), NMDS/ANOSIM and PERMANOVA analyses both showed association between
482 the vaginal microbiome and vaginal pH (Table 1, S5-S7 Files). Using BN analysis, we were able
483 to reveal that vaginal pH was influenced by 4 bacteria: *Atopobium*, *Dialister, Gardnerella* and
484 *Peptoniphilus* (Fig. 5). These bacteria are all obligate or facultative anaerobes and the presence
485 of each of these bacteria in the vaginal microbiome increased the likelihood of a vaginal pH >
486 4.5. These four bacteria have been described as key members of a vaginal microbiome
487 community type characterized by a diversity of non-*Lactobacilli* bacteria that is found more
488 commonly in women with a high vaginal pH. It has been hypothesized that the higher vaginal pH
489 in such women is due to a comparatively low number of lactic acid producing bacteria (8), and
490 the results of this study support this hypothesis. As with previous research, however, the
491 presence of these bacteria and the subsequent increase in vaginal pH were reported in healthy
492 women without clinical BV.

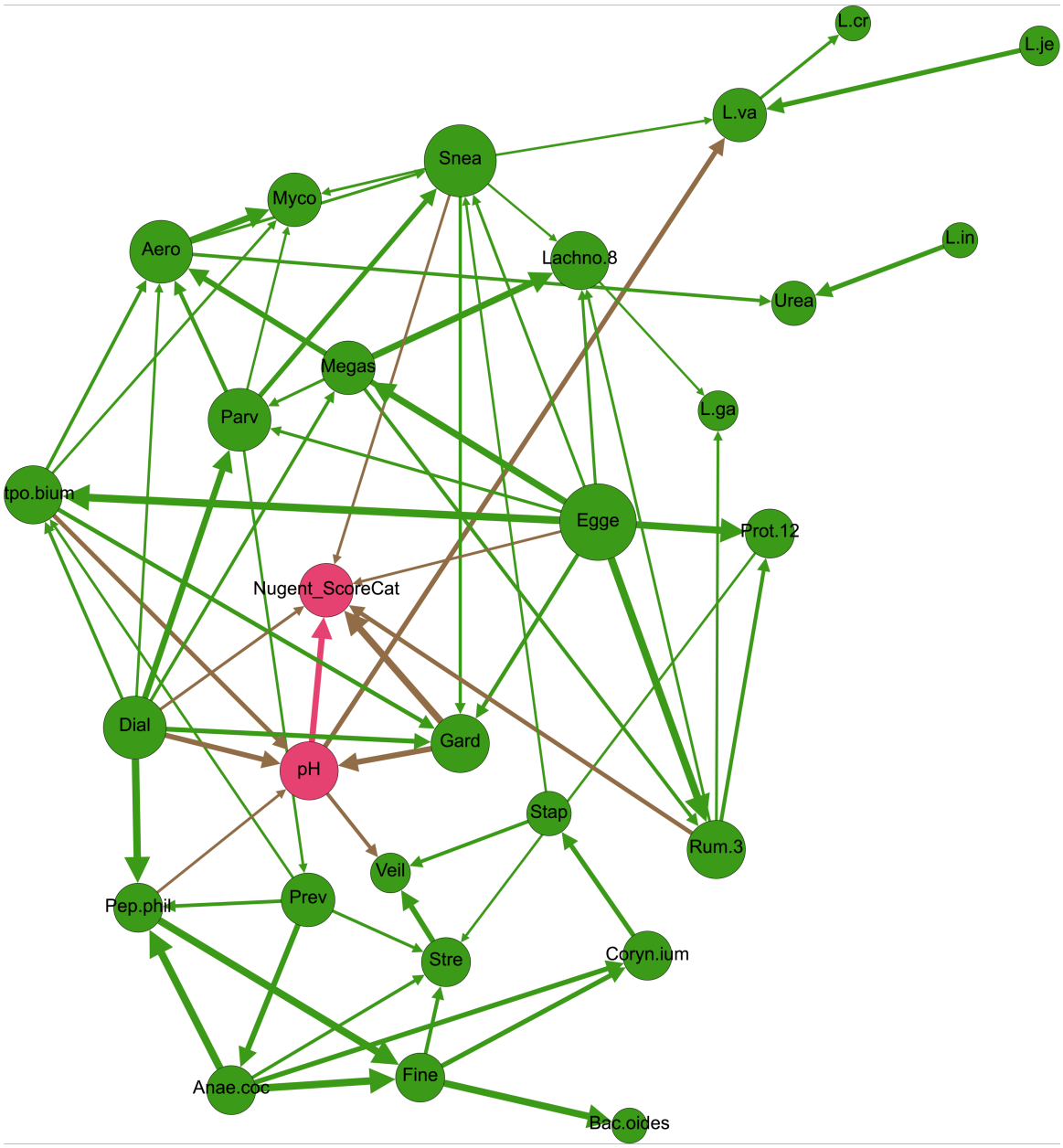
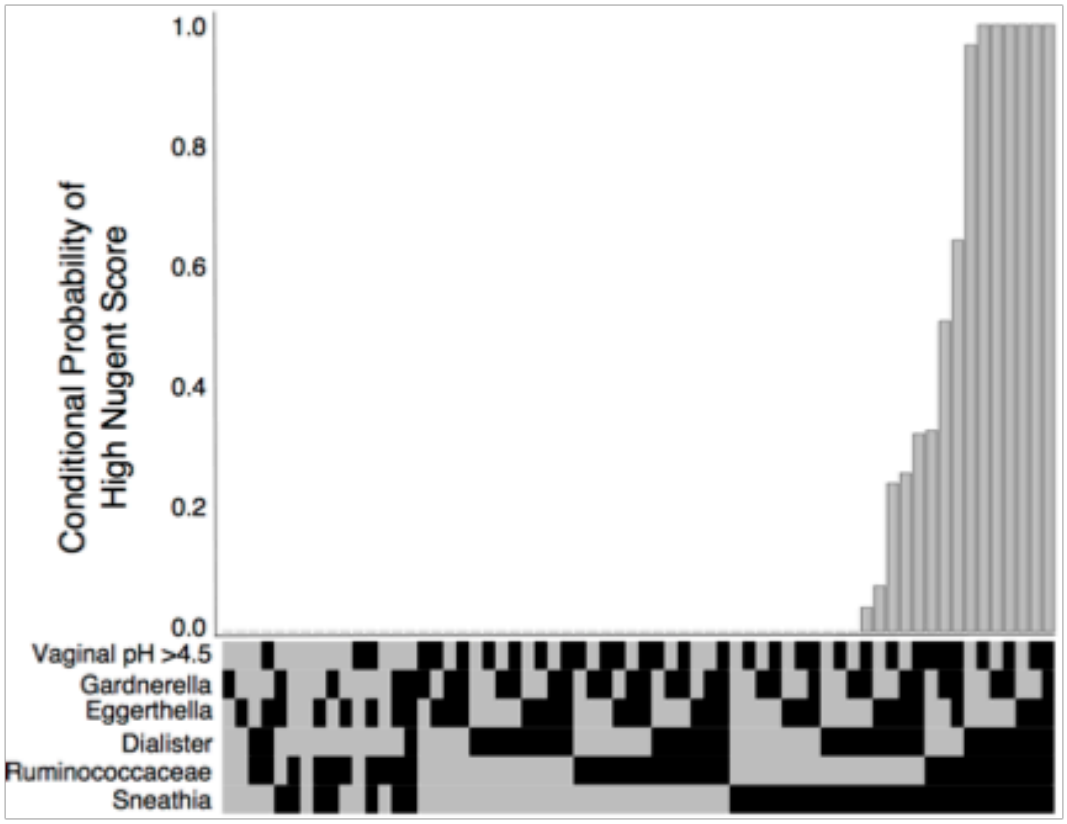


Figure 5. Network depicting vaginal pH and Nugent Score (pink), and bacteria (green). Size of circle is proportional to degree (i.e., number of incoming and outgoing edges), and edge thickness is proportional to bootstrap support. Node label abbreviations can be found in S4 File.

497 The Nugent score is a diagnostic test used primarily in research studies of BV (13). It is rarely
498 used in practice because it involves time-consuming microscopic examination of bacterial
499 morphology from vaginal swabs. The Nugent score is based on the relative presence of Gram-
500 positive versus Gram-negative straight and curved rods, and therefore it is unsurprising that
501 NMDS/ANOSIM and PERMANOVA testing both uncovered associations between the vaginal
502 microbiome and the Nugent score (Table 1, S5-S7 Files). Using BN analysis, we were able to
503 assess this association more closely in order to uncover which specific microbiome members
504 were driving differences in the Nugent score. We found that the Nugent score for the 285
505 women in this dataset was influenced by bacteria from two Gram-positive bacteria (*Eggerthella*
506 and *Ruminococcaceae*), two Gram-negative bacteria (*Dialister* and *Sneathia*) and one Gram-
507 variable bacterium (*Gardnerella*) group. Vaginal pH was also connected to Nugent score;
508 women with a pH of >4.5 were much more likely to have a high Nugent score, a finding
509 consistent with previous literature (8,62). Amongst the 64 combinations of parent variable states
510 that directly influenced the Nugent Score, there were 6 combinations that conferred 100%
511 conditional probability of a high Nugent score; all 6 combinations included *Dialister*,
512 *Ruminococcaceae* and *Sneathia*, which have been found in higher abundance in women with a
513 high Nugent score (8, Fig. 6). Overall, these results confirm previous findings concerning the
514 robust association between certain vaginal microbiome members and vaginal pH and Nugent
515 score. However, the women in this study were BV-negative, and therefore the presence of high
516 vaginal pH, a high Nugent score and/or microbes associated with these indicators does not
517 always correlate with clinical BV (8). It is worth noting that the *Lactobacillus* species and the
518 Nugent Score were only conditionally independent given the presence of vaginal pH and
519 *Sneathia*—which were directly associated with *L. vaginalis* and indirectly with both *L. jensenii*

520 and *L. crispatus*—and *Ruminococcacea*, which shared a direct edge with *L. gasseri*. This
521 indicates a weaker association between the Nugent Score and the *Lactobacilli*, which was
522 masked by the presence of stronger associations, e.g., with vaginal pH. Removing some of these
523 directly-linked nodes from the network resulted in a direct association of the Nugent score with
524 some *Lactobacillus* nodes (data not shown).



525
526 **Figure 6. Conditional probability of a high Nugent Score (y-axis) for all possible**
527 **combinations of direct parent variables (x-axis) (binary heatmap, black = present, gray =**
528 **absent).**

529 **Associations between the microbiome and risk symptoms other than**
530 **vaginal pH**

531 While the women in this study did not have clinical BV, 108 of them did report experiencing at
532 least one risk symptom associated with BV within the 60 days prior to sample collection
533 (108/285, 38.9%). Vaginal yeast infections were reported by 19 of the 285 respondents during
534 the same time period. Most risk symptoms clustered together within the BN, with edges
535 connecting vaginal discharge to both vaginal irritation and vaginal itch; a fishy/musty vaginal
536 odor to staining of the underwear; any vaginal odor to vaginal itch; vaginal itch to vaginal
537 infection; and vaginal infection to yeast infection specifically. Staining of the underwear was
538 also directly connected to uro-abdominal pain in the preceding 60 days, which in turn was
539 connected to self-treatment of a vaginal infection. While the Nugent score and vaginal pH were
540 not connected to these risk symptoms, there were direct connections with specific microbial taxa.
541 Presence of *Lactobacillus jensenii* increased the likelihood of a woman reporting uro-abdominal
542 pain within the preceding 60 days, from a probability of 3.6% to 10.1% conditional on absence
543 or presence of *L. jensenii*, respectively. Interestingly, PERMANOVA analysis showed an
544 association between underwear staining in the previous 60 days and the vaginal microbiome
545 (Table 1); the results of BN analysis suggest that uro-abdominal pain and *L. jensenii* may be the
546 specific factors that link underwear staining and the microbiome. Use of pantyliners during
547 menstruation was also associated with *L. jensenii*, with the conditional probability of *L. jensenii*
548 being present in the vaginal microbiome increasing from 37.4% to 60.7% with use of panty
549 liners during menstruation. While the directionality of this relationship is not immediately
550 intuitive, previous work has shown that use of emollient pads changes the vaginal epithelium and
551 that some women's vaginal microflora does shift with pad versus tampon use (63,64), although
552 the evidence on this is mixed (65). NMDS/ANOSIM testing indicated that tampon use was
553 associated with the vaginal microbiome, lending further support to the hypothesis that menstrual

554 habits could impact the microbiome (S5 and S6 Files). Alternatively, the support for edge
555 directionality in the Bayesian network was based on majority rule and therefore directionality of
556 arcs could potentially be reversed; in such a scenario, presence of *L. jensenii* could increase the
557 likelihood of uro-abdominal pain, and thus underwear staining and pantyliner use. Furthermore,
558 ethnicity was a potential confounder in the relationship between the microbiome and menstrual
559 habits, as the network showed an influence of ethnicity on pad versus tampon use, which in turn
560 directly affected pantyliner use (Fig. 4). The relationships between uro-abdominal pain,
561 underwear staining, pantyliner use and *L. jensenii* deserve closer study given the importance of
562 this microbe in vaginal health.

563 Other connections between bacteria and risk symptoms of itch, irritation, odor, staining, pain or
564 yeast infection included use of a pad outside of menstruation increasing the conditional
565 likelihood of a woman harboring *Proteobacteria* from 2.8% to 6.7%, as extra-menstrual pad use
566 could be considered a proxy for vaginal discharge, staining or odor. Because *Proteobacteria* is
567 such a diverse phylum that includes pathogens and non-pathogens, it is difficult to formulate
568 hypotheses concerning these relationships. *Corynebacterium* seemed to play an important role in
569 symptomology and sexual behaviors, as it was directly connected to vaginal odor and number of
570 sexual partners in the 60 days preceding the survey. Women with this bacterium in their vaginal
571 flora were more likely to have reported both no or multiple sexual partners, and were less likely
572 to have reported a single sexual partner, while presence of *Corynebacterium* increased the risk of
573 vaginal odor in the 60 days preceding the survey. The association of *Corynebacterium* with
574 vaginal pathology is not widely described in the literature, although it has been recognized as an
575 important vaginal community member for several decades (66). More recently, studies indicate
576 that this bacterium is more prevalent in women with spinal cord injuries, women with treatment-

577 refractory BV, and women with urgency urinary incontinence (67–69). Depending on the
578 species, *Corynebacterium* can be fermentative, and many species produce amino acids, which
579 can alter pH and the metabolite profile of the surrounding environment, thus providing a
580 potential mechanism of *Corynebacterium's* role in vaginal odor and pathology. These
581 associations were not and would not have been easily observed using NMDS/ANOSIM and
582 PERMANOVA.

583 Presence of *Bacteroides* had a protective effect against vaginal discharge; the probability of
584 reporting discharge given the presence of *Bacteroides* was 2.0%, compared to 28.9% given
585 absence of these bacteria within the vaginal microbiome. This runs counter to current notions
586 about *Bacteroides*, an anerobic bacteria which is considered a commensal of the gastrointestinal
587 tract and has been implicated in BV (70,71). However, not all vaginal discharge should be
588 considered abnormal (72,73), and therefore it would not be a contradiction for *Bacteroides* to be
589 associated with a decreased prevalence of "healthy" vaginal discharge (such as cervical mucous
590 released during ovulation) and an increased prevalence of clinical BV. Distinctions such as these
591 will become increasingly important as the scientific community attempts to attain a more
592 nuanced understanding of connections between the microbiome and true clinical disease.

593 Microbiome-Behavior Connections

594 Sexual behavior, sanitary practices and microbiome composition have all been associated with
595 BV as separate factors. However, little is known about how a woman's sexual and sanitary
596 habits may influence the microbiome, or vice versa; and how this interplay may influence the
597 likelihood of developing BV. The network resulting from this dataset showed several direct
598 connections between the study participants' habits and the microbiome. Whether or not women

599 had ever had a pap smear was directly related to presence/absence of *Ureaplasma*. Women who
600 reported having had a pap demonstrated a 34.1% likelihood of harboring *Ureaplasma*, compared
601 to 18.0% among women who had never had a pap smear. Having had a Pap test was the only
602 healthcare access question included in the Bayesian network that could be considered a proxy for
603 socioeconomic status (48).

604 Microbiome-Birth Control Interactions

605 Use of certain types of birth control could influence the microbiome by modulating hormone
606 levels or by introducing chemical or physical barriers into the vaginal environment. Use of an
607 intrauterine device (IUD) directly influenced the presence/absence of two different bacteria,
608 namely *Lactobacillus gasseri* and *Sneathia*; in both cases, use of an IUD decreased the
609 likelihood of finding the bacterium within the vaginal microbiome. Previous research on the
610 effects of IUD use on vaginal microflora has been mixed, with both copper and hormone IUDs
611 associated with higher levels of *Candida* fungi (74) but no associations with any bacterial
612 alterations (74–76). Given the location of IUD use as a relatively highly-connected intercategory
613 node in the Bayesian network, as well as its potential association with nonspecific vaginitis and
614 recurrent BV (32,38,68), the associations with *L. gasseri* and *Sneathia* warrant follow-up study.

615 Women who reported ever using any form of birth control were more likely to have
616 *Lactobacillus crispatus* within their microbiome compared to women who had never used birth
617 control (conditional probability of 56.6% versus 48.1%, respectively). Women who reported
618 male condom use in the 60 days preceding the survey were less likely to harbor *Lachnospiraceae*
619 than women who did not (conditional probability of 1.3% versus 9.9%). Interestingly,
620 PERMANOVA testing also showed a connection between male condom use in the previous 60

621 days and the overall composition of the vaginal microbiome (Table 1); the BN analysis allowed
622 us to pinpoint this relationship to *Lachnospiraceae* specifically. As with other associations, there
623 were some directed edges that did not immediately fit the paradigm for host-microbe
624 interactions. For example, IUD use and abstinence were influenced by status of *Parvimonas* and
625 *L. vaginalis* within the vagina, respectively. It is difficult to hypothesize a mechanism by which
626 vaginal bacteria could influence a woman's birth control decisions, even though recent studies
627 have challenged the paradigm that host-microbiome interactions are largely uni-directional by
628 suggesting that microbiome composition can directly influence host behavior (77–79). This
629 highlights the importance of remembering that the directionality of edges in the BN are not
630 definitive.

631 Microbiome-Microbiome Interactions

632 Within the entire network, bacterial nodes fell largely within 2 modularity-based groupings,
633 with some exceptions (Fig. 4). Upon isolation of bacterial nodes in a separate graph and
634 subsequent modularity optimization, four communities of bacteria became evident (Fig. 7).
635 These groupings did not seem to segregate by either Gram stain, cellular metabolism, or any
636 other known characteristics. These groupings could be driven by as-yet undiscovered microbial
637 relationships within the vaginal environment and/or by unmeasured, non-microbiome factors in
638 this dataset.

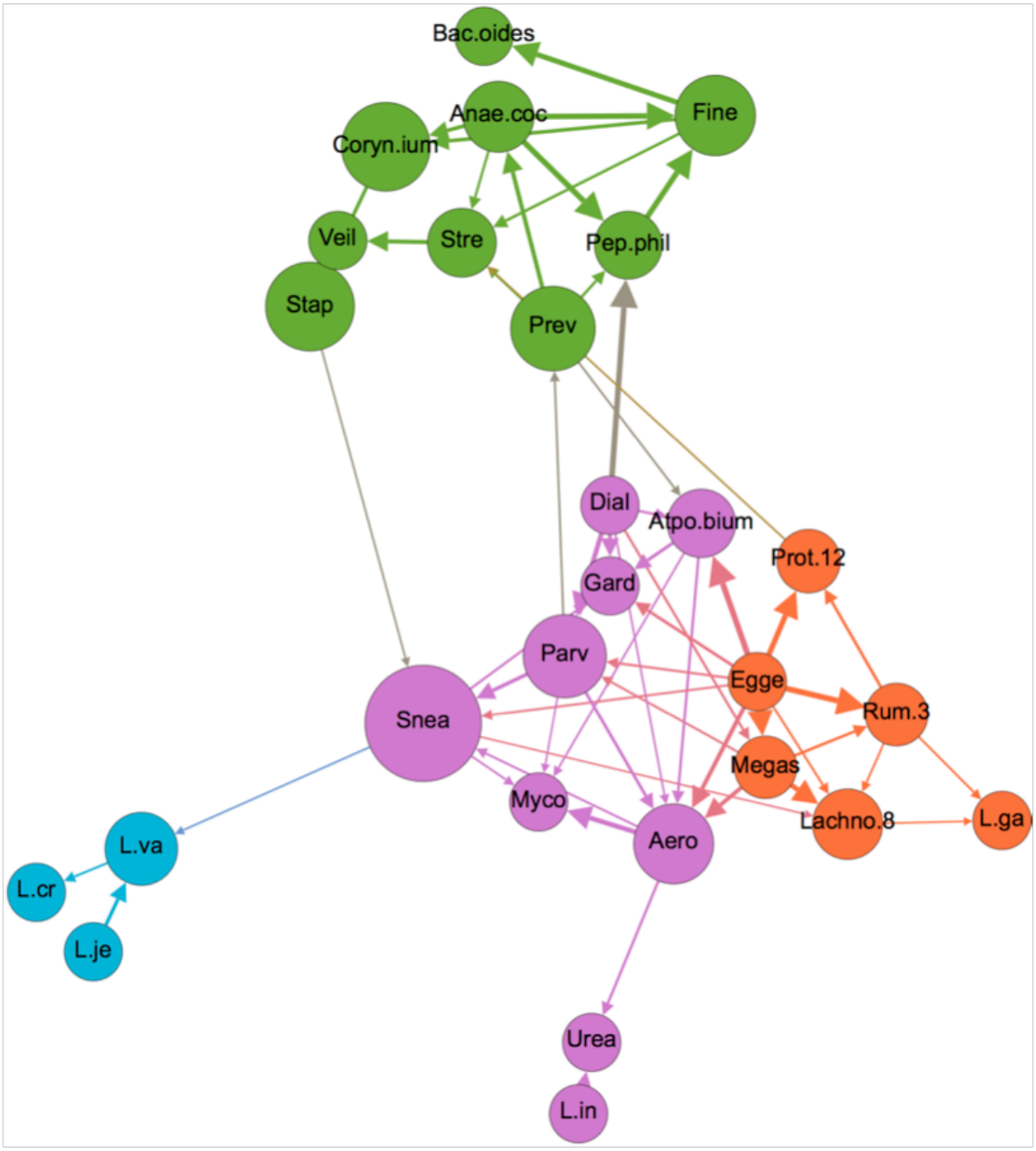


Figure 7. Network depicting only microbiome nodes, colored by modularity-based community type, as determined using a modularity optimizing algorithm (41). Node size is proportional to betweenness centrality, and edge thickness is proportional to bootstrap support. Node label abbreviations can be found in S4 File.

644 With the exception of *L. iners* and *L. gasseri*, the *Lactobacilli* formed a tight cluster to
645 themselves. The fact that *L. iners* segregated separately from the other *Lactobacilli* is interesting
646 given that this organism differs substantially in terms of genomic and metabolic characteristics,
647 and its role in vaginal health has been debated despite its very high prevalence (47,80). One clue
648 in this debate could be the tight association between *L. iners* and *Ureaplasma*, the latter of which
649 has been implicated in several female reproductive conditions including chorioamnionitis leading
650 to premature delivery and pelvic inflammatory disease (81,82). Presence of *L. iners* increased the
651 likelihood of *Ureaplasma* from 19.4% to 36.3%, a significant increase that could have health
652 implications and therefore deserves follow-up study.

653

654 # Conclusions

655 Using widely accessible software, we have demonstrated the utility of applying a Bayesian
656 Network approach to a multi-dimensional microbiome dataset. Using this approach, we have
657 demonstrated associations between women's sexual and menstrual habits, demographics, vaginal
658 microbiome composition and symptoms and diagnostics of BV. Many of these associations
659 suggested intriguing relationships, indicating that the BN approach is able to highlight important
660 associations within complex datasets, which can then be used for hypothesis generation. While
661 follow-up studies are needed to investigate the significance of these novel associations, the
662 validity of the associations was buttressed by the presence of many well-documented and self-
663 evident connections within the overall BN. For instance, our BN confirmed the importance of
664 vaginal pH and *Gardnerella* as influencers on the Nugent Score (Fig. 4). In addition, we found a

665 very strong association between previous pregnancy, vaginal birth and tubal ligation. Indeed, all
666 17 participants who reported tubal ligation had previously been pregnant. The most recent data
667 available suggest that nearly half of tubal ligations were performed in the postpartum period and
668 that younger women were more likely to have tubal ligation performed post-partum (83). As
669 with the BN produced here, this suggests a tight association between tubal ligation and previous
670 pregnancy. Another confirmatory association showed that women who reported vaginal itch in
671 the 60 days preceding the survey were much more likely to also have reported a yeast infection
672 in the same time period. Vaginal itch has been closely associated with vulvovaginitis candidiasis
673 (yeast infection) (84). While relationships such as these do not provide novel insight, they can be
674 used as a "sanity check" in order to gain confidence in the structure of the BN to then move on to
675 analysis of less well-understood network connections.

676 An advantage of the BN approach to datasets such as the one presented here is that BN
677 algorithms are now easily accessible through well-documented and well-supported packages
678 such as bnlearn (29), which has been integrated with the packages "parallel" (which is included
679 in R-core) and "snow" (85) to support multi-threading for construction of large networks, and
680 which now supports both categorical and continuous variables. Numerous GUI-based packages
681 exist to support network visualization once the BN has been constructed (42,86), allowing for
682 interactive graph exploration in an intuitive interface. Given the increasing use of multi-level,
683 multivariate microbiome datasets, this accessibility should pave the way for more scientists to
684 implement the BN analysis for such data. The need for such an approach has been demonstrated
685 in the complexity of the relationships that we found within and between the vaginal microbiome,
686 women's sexual and menstrual habits, demographics and symptoms/diagnostics. While non-
687 network approaches such as PERMANOVA and ordination/clustering can be used to uncover

688 associations between individual metadata variables and the microbiome as a whole (as we
689 demonstrate in this work), there are important limitations to these methods: first, these methods
690 collapse the microbiome data into a single datapoint projected into two or a maximum of three
691 dimensions, and therefore preclude assessment of intra-microbiome interactions, as well as
692 interactions between metadata variables and specific microbiome members; second, these
693 methods usually approach statistical testing on a variable-by-variable basis, and are thus ill-
694 suited to understanding complex interactions between subsets of variables and the microbiome;
695 and third, these approaches fail to show intra-metadata interactionswithin an analysis. BN
696 analysis overcomes these limitations by testing for all possible associations between every node
697 within the data – thus uncovering complex interactions within hierarchical, multidimensional
698 datasets. Such datasets increasingly typify microbiome studies, and the need for rich and
699 standardized metadata to support microbiome studies has been recently noted and implemented
700 (87,88); such initiatives would be well-complemented by the data-flexible and intuitive nature of
701 BN's. It is important to note, however, that the incorporation of more metadata variables (nodes)
702 into BN analyses will necessitate larger sample sizes in order to support BN construction and
703 inference. Similarly, the data distributions currently accommodated by available BN analysis
704 packages for continuous variables are limited to the Normal (Gaussian) distribution. This
705 constraint results in the need to discretize most microbiome data, which are commonly
706 characterized by zero-inflated, skewed counts that do not conform to a Gaussian distribution. As
707 was the case in this report, the need to dichotomize may cause loss of information, resulting in
708 less-than-optimally refined inference. Fortunately, BN analysis methods form an active area of
709 research, and extensions to accommodate distributions other than the Gaussian are currently

being developed (89). As these efforts evolve, the BN approach is likely to become more flexible, and thus even better-suited to the unique characteristics of microbiome datasets.

Acknowledgements

We would like to thank the editor and three anonymous reviewers for there useful comments.

References

1. Donders GG, Bosmans E, Dekeersmaecker A, Vereecken A, Van Bulck B, Spitz B. Pathogenesis of abnormal vaginal bacterial flora. Am J Obstet Gynecol. 2000 Apr;182(4):872–8.

2. Pybus V, Onderdonk AB. Microbial interactions in the vaginal ecosystem, with emphasis on the pathogenesis of bacterial vaginosis. Microbes Infect Inst Pasteur. 1999 Apr;1(4):285–92.

3. Martin HL, Richardson BA, Nyange PM, Lavreys L, Hillier SL, Chohan B, et al. Vaginal lactobacilli, microbial flora, and risk of human immunodeficiency virus type 1 and sexually transmitted disease acquisition. J Infect Dis. 1999 Dec;180(6):1863–8.

4. Allsworth JE, Peipert JF. Prevalence of bacterial vaginosis: 2001-2004 National Health and Nutrition Examination Survey data. Obstet Gynecol. 2007 Jan;109(1):114–20.

5. Zhou X, Brotman RM, Gajer P, Abdo Z, Schüette U, Ma S, et al. Recent advances in understanding the microbiology of the female reproductive tract and the causes of premature birth. Infect Dis Obstet Gynecol. 2010;2010:737425.

6. Bautista CT, Wurapa E, Sateren WB, Morris S, Hollingsworth B, Sanchez JL. Bacterial vaginosis: a synthesis of the literature on etiology, prevalence, risk factors, and relationship with chlamydia and gonorrhea infections. Mil Med Res. 2016;3:4.

7. Eschenbach DA, Davick PR, Williams BL, Klebanoff SJ, Young-Smith K, Critchlow CM, et al. Prevalence of hydrogen peroxide-producing Lactobacillus species in normal women and women with bacterial vaginosis. J Clin Microbiol. 1989 Feb;27(2):251–6.

8. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, et al. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci. 2011 Mar 15;108(Supplement 1):4680–7.

738 9. Fredricks DN, Fiedler TL, Marrazzo JM. Molecular Identification of Bacteria Associated
739 with Bacterial Vaginosis. N Engl J Med. 2005 Nov 3;353(18):1899–911.

740 10. Ferris MJ, Norori J, Zozaya-Hinchliffe M, Martin DH. Cultivation-Independent Analysis
741 of Changes in Bacterial Vaginosis Flora Following Metronidazole Treatment. J Clin
742 Microbiol. 2007 Mar;45(3):1016–8.

743 11. Lamont RF, Sobel JD, Akins RA, Hassan SS, Chaiworapongsa T, Kusanovic JP, et al. The
744 vaginal microbiome: new information about genital tract flora using molecular based
745 techniques. BJOG Int J Obstet Gynaecol. 2011 Apr;118(5):533–49.

746 12. Livengood CH. Bacterial Vaginosis: An Overview for 2009. Rev Obstet Gynecol.
747 2009;2(1):28–37.

748 13. Nugent RP, Krohn MA, Hillier SL. Reliability of diagnosing bacterial vaginosis is
749 improved by a standardized method of gram stain interpretation. J Clin Microbiol.
750 1991 Feb;29(2):297–301.

751 14. Ma B, Forney LJ, Ravel J. The vaginal microbiome: rethinking health and diseases.
752 Annu Rev Microbiol. 2012;66:371–89.

753 15. Beigi RH, Wiesenfeld HC, Hillier SL, Straw T, Krohn MA. Factors associated with
754 absence of H2O2-producing Lactobacillus among women with bacterial vaginosis. J
755 Infect Dis. 2005 Mar 15;191(6):924–9.

756 16. Brotman RM, Klebanoff MA, Nansel TR, Andrews WW, Schwebke JR, Zhang J, et al. A
757 Longitudinal Study of Vaginal Douching and Bacterial Vaginosis—A Marginal
758 Structural Modeling Analysis. Am J Epidemiol. 2008 Jul 15;168(2):188–96.

759 17. Hellberg D, Nilsson S, Mårdh PA. Bacterial vaginosis and smoking. Int J STD AIDS. 2000
760 Sep;11(9):603–6.

761 18. Hay PE, Ugwumadu A, Chowns J. Sex, thrush and bacterial vaginosis. Int J STD AIDS.
762 1997 Oct;8(10):603–8.

763 19. Ness RB, Hillier S, Richter HE, Soper DE, Stamm C, Bass DC, et al. Can known risk
764 factors explain racial differences in the occurrence of bacterial vaginosis? J Natl Med
765 Assoc. 2003 Mar;95(3):201–12.

766 20. Zhou X, Hansmann MA, Davis CC, Suzuki H, Brown CJ, Schütte U, et al. The vaginal
767 bacterial communities of Japanese women resemble those of women in other racial
768 groups. FEMS Immunol Med Microbiol. 2010 Mar;58(2):169–81.

769 21. Warton DI, Wright ST, Wang Y. Distance-based multivariate analyses confound
770 location and dispersion effects. Methods Ecol Evol. 2012 Feb 1;3(1):89–101.

771 22. Gloor GB, Reid G. Compositional analysis: a valid approach to analyze microbiome
772 high-throughput sequencing data. Can J Microbiol. 2016 Apr 12;62(8):692–703.

773 23. Xia Y, Sun J. Hypothesis testing and statistical analysis of microbiome. Genes Dis
774 [Internet]. 2017 Jun 23 [cited 2017 Aug 27]; Available from:
775 http://www.sciencedirect.com/science/article/pii/S2352304217300351

776 24. Shankar J. Insights into study design and statistical analyses in translational
777 microbiome studies. Ann Transl Med [Internet]. 2017 Jun [cited 2017 Aug 27];5(12).
778 Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5497086/

779 25. Ramoni RB, Himes BE, Sale MM, Furie KL, Ramoni MF. Predictive Genomics of
780 Cardioembolic Stroke. Stroke J Cereb Circ. 2009 Mar;40(3 Suppl):S67–70.

781 26. Gevaert O, Smet FD, Timmerman D, Moreau Y, Moor BD. Predicting the prognosis of
782 breast cancer by integrating clinical and microarray data with Bayesian networks.
783 Bioinformatics. 2006 Jul 15;22(14):e184–90.

784 27. Cruz-Ramírez N, Acosta-Mesa HG, Carrillo-Calvet H, Nava-Fernández LA, Barrientos-
785 Martínez RE. Diagnosis of breast cancer using Bayesian networks: a case study.
786 Comput Biol Med. 2007 Nov;37(11):1553–64.

787 28. Gámez JA, Mateo JL, Puerta JM. Learning Bayesian networks by hill climbing: efficient
788 methods based on progressive restriction of the neighborhood. Data Min Knowl
789 Discov. 2010 May 11;22(1-2):106–48.

790 29. Scutari M. Learning Bayesian Networks with the bnlearn R Package. J Stat Softw.
791 2010;35(3):1–22.

792 30. Højsgaard S. Graphical Independence Networks with the gRain Package for R. J Stat
793 Softw. 2012;46(10):1–26.

794 31. Uusitalo L. Advantages and challenges of Bayesian networks in environmental
795 modelling. Ecol Model. 2007 May 10;203(3–4):312–8.

796 32. Amsel R, Totten PA, Spiegel CA, Chen KC, Eschenbach D, Holmes KK. Nonspecific
797 vaginitis. Diagnostic criteria and microbial and epidemiologic associations. Am J Med.
798 1983 Jan;74(1):14–22.

799 33. Gajer P, Brotman RM, Bai G, Sakamoto J, Schütte UME, Zhong X, et al. Temporal
800 Dynamics of the Human Vaginal Microbiota. Sci Transl Med. 2012 May
801 2;4(132):132ra52–132ra52.

802 34. Anderson MJ. A new method for non‐parametric multivariate analysis of variance.
803 Austral Ecol. 2001 Feb 1;26(1):32–46.

35. Paulson JN, Stine OC, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. Nat Methods. 2013 Dec;10(12):1200–2.

36. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. Appl Environ Microbiol. 2007 Aug 15;73(16):5261–7.

37. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2016. Available from: https://www.R-project.org/

38. Li X-D, Wang C-C, Zhang X-J, Gao G-P, Tong F, Li X, et al. Risk factors for bacterial vaginosis: results from a cross-sectional study having a sample of 53,652 women. Eur J Clin Microbiol Infect Dis Off Publ Eur Soc Clin Microbiol. 2014 Sep;33(9):1525–32.

39. Steck H. Learning the Bayesian Network Structure: Dirichlet Prior versus Data. Proc 24th Conf Uncertain Artif Intell. 2008;511–8.

40. Nwankwo TO, Aniebue UU, Umeh UA. Syndromic Diagnosis in Evaluation of Women with Symptoms of Vaginitis. Curr Infect Dis Rep. 2017 Jan 1;19(1):3.

41. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech Theory Exp. 2008 Oct 9;2008(10):P10008.

42. Bastian M, Heymann S, Jacomy M. Gephi: An Open Source Software for Exploring and Manipulating Networks. In San Jose, California, USA; 2009.

43. Legendre P, Gallagher ED. Ecologically meaningful transformations for ordination of species data. Oecologia. 2001 Oct;129(2):271–80.

44. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. vegan: Community Ecology Package [Internet]. 2014. Available from: http://CRAN.R-project.org/package=vegan

45. Clarke KR. Non-parametric multivariate analyses of changes in community structure. Aust J Ecol. 1993 Mar 1;18(1):117–43.

46. Gelman A, Hill J. Data analysis using regression and multivariate/hierarchical models. New York, NY: Cambridge University Press; 2007. 625 p. (Analytical Methods for Social Research).

47. Srinivasan S, Hoffman NG, Morgan MT, Matsen FA, Fiedler TL, Hall RW, et al. Bacterial Communities in Women with Bacterial Vaginosis: High Resolution Phylogenetic Analyses Reveal Relationships of Microbiota to Clinical Criteria. PLOS ONE. 2012 Jun 18;7(6):e37818.

837 48. Monnat SM. Race/Ethnicity and the Socioeconomic Status Gradient in Women's Cancer
838 Screening Utilization: A Case of Diminishing Returns? J Health Care Poor Underserved.
839 2014 Feb;25(1):332–56.

840 49. Sauer AG, Jemal A, Simard EP, Fedewa SA. Differential uptake of recent Papanicolaou
841 testing by HPV vaccination status among young women in the United States, 2008–
842 2013. Cancer Epidemiol. 2015 Aug;39(4):650–5.

843 50. Romo LF, Berenson AB. Tampon use in adolescence: differences among European
844 American, African American and Latina women in practices, concerns, and barriers. J
845 Pediatr Adolesc Gynecol. 2012 Oct;25(5):328–33.

846 51. Anderson SE, Dallal GE, Must A. Relative weight and race influence average age at
847 menarche: results from two nationally representative surveys of US girls studied 25
848 years apart. Pediatrics. 2003 Apr;111(4 Pt 1):844–50.

849 52. Chumlea WC, Schubert CM, Roche AF, Kulin HE, Lee PA, Himes JH, et al. Age at
850 menarche and racial comparisons in US girls. Pediatrics. 2003 Jan;111(1):110–3.

851 53. Nunn KL, Forney LJ. Unraveling the Dynamics of the Human Vaginal Microbiome. Yale
852 J Biol Med. 2016 Sep 30;89(3):331–7.

853 54. Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Nikita L, et al. The composition
854 and stability of the vaginal microbiota of normal pregnant women is different from
855 that of non-pregnant women. Microbiome. 2014 Feb 3;2(1):4.

856 55. Horstman AM, Dillon EL, Urban RJ, Sheffield-Moore M. The Role of Androgens and
857 Estrogens on Healthy Aging and Longevity. J Gerontol A Biol Sci Med Sci. 2012
858 Nov;67(11):1140–52.

859 56. Tibaldi C, Cappello N, Latino MA, Polarolo G, Masuelli G, Cavallo F, et al. Maternal risk
860 factors for abnormal vaginal flora during pregnancy. Int J Gynaecol Obstet Off Organ
861 Int Fed Gynaecol Obstet. 2016 Apr;133(1):89–93.

862 57. Carey J, Blackwelder W, Nugent R, Matteson M, Rao A, Eschenbach D, et al. Antepartum
863 Cultures for Ureaplasma-Urealyticum Are Not Useful in Predicting Pregnancy
864 Outcome. Am J Obstet Gynecol. 1991 Mar;164(3):728–33.

865 58. Goldenberg RL, Hauth JC, Andrews WW. Intrauterine Infection and Preterm Delivery.
866 N Engl J Med. 2000 May 18;342(20):1500–7.

867 59. Freedman DS, Khan LK, Serdula MK, Dietz WH, Srinivasan SR, Berenson GS. Relation of
868 age at menarche to race, time period, and anthropometric dimensions: the Bogalusa
869 Heart Study. Pediatrics. 2002 Oct;110(4):e43.

870 60. Wu T, Mendola P, Buck GM. Ethnic differences in the presence of secondary sex
871 characteristics and menarche among US girls: the Third National Health and Nutrition
872 Examination Survey, 1988-1994. Pediatrics. 2002 Oct;110(4):752–7.

873 61. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, et al. Vaginal
874 microbiome of reproductive-age women. Proc Natl Acad Sci. 2011 Mar
875 15;108(Supplement 1):4680–7.

876 62. Hoffman MK, Bellad MB, Charantimath US, Kavi A, Nagmoti JM, Nagmoti MB, et al. A
877 Comparison of Colorimetric Assessment of Vaginal pH with Nugent Score for the
878 Detection of Bacterial Vaginosis. Infect Dis Obstet Gynecol. 2017 Feb
879 15;2017:e1040984.

880 63. Farage MA, Miller KW, Azuka CE, Wang B, Lankhof S, Henn D, et al. Skin benefits with a
881 novel emollient-treated menstrual pad. Rev Recent Clin Trials. 2013 Mar;8(1):29–35.

882 64. Hickey RJ, Abdo Z, Zhou X, Nemeth K, Hansmann M, Osborn TW, et al. Effects of
883 tampons and menses on the composition and diversity of vaginal microbial
884 communities over time. BJOG Int J Obstet Gynaecol. 2013 May;120(6):695–704;
885 discussion 704–6.

886 65. Klebanoff MA, Nansel TR, Brotman RM, Zhang J, Yu K-F, Schwebke JR, et al. Personal
887 hygienic behaviors and bacterial vaginosis. Sex Transm Dis. 2010 Feb;37(2):94–9.

888 66. McCormack WM, Hayes CH, Rosner B, Evrard JR, Crockett VA, Alpert S, et al. Vaginal
889 colonization with Corynebacterium vaginale (Haemophilus vaginalis). J Infect Dis.
890 1977 Dec;136(6):740–5.

891 67. Pires CVG, Linhares IM, Serzedello F, Fukazawa EI, Baracat EC, Witkin SS. Alterations
892 in the Genital Microbiota in Women With Spinal Cord Injury. Obstet Gynecol. 2016
893 Feb;127(2):273–8.

894 68. Wang B, Xiao BB, Shang CG, Wang K, Na RS, Nu XX, et al. Molecular analysis of the
895 relationship between specific vaginal bacteria and bacterial vaginosis metronidazole
896 therapy failure. Eur J Clin Microbiol Infect Dis Off Publ Eur Soc Clin Microbiol. 2014
897 Oct;33(10):1749–56.

898 69. Pearce MM, Hilt EE, Rosenfeld AB, Zilliox MJ, Thomas-White K, Fok C, et al. The female
899 urinary microbiome: a comparison of women with and without urgency urinary
900 incontinence. mBio. 2014 Jul 8;5(4):e01283–01214.

901 70. Verhelst R, Verstraelen H, Claeys G, Verschraegen G, Delanghe J, Van Simaey L, et al.
902 Cloning of 16S rRNA genes amplified from normal and disturbed vaginal microflora
903 suggests a strong association between Atopobium vaginae, Gardnerella vaginalis and
904 bacterial vaginosis. BMC Microbiol. 2004 Apr 21;4:16.

905 71. Machado D, Castro J, Palmeira-de-Oliveira A, Martinez-de-Oliveira J, Cerca N. Bacterial
906 Vaginosis Biofilms: Challenges to Current Therapies and Emerging Solutions. Front
907 Microbiol [Internet]. 2016 [cited 2017 Feb 10];6. Available from:
908 http://journal.frontiersin.org/article/10.3389/fmicb.2015.01528/abstract

909 72. Zinaman MJ. Using cervical mucus and other easily observed biomarkers to identify
910 ovulation in prospective pregnancy trials. Paediatr Perinat Epidemiol. 2006 Nov;20
911 Suppl 1:26–9.

912 73. Powell AM, Nyirjesy P. New Perspectives on the Normal Vagina and Noninfectious
913 Causes of Discharge. Clin Obstet Gynecol. 2015 Sep;58(3):453–63.

914 74. Erol O, Simavlı S, Derbent AU, Ayrım A, Kafalı H. The impact of copper-containing and
915 levonorgestrel-releasing intrauterine contraceptives on cervicovaginal cytology and
916 microbiological flora: a prospective study. Eur J Contracept Reprod Health Care Off J
917 Eur Soc Contracept. 2014 Jun;19(3):187–93.

918 75. Jacobson JC, Turok DK, Dermish AI, Nygaard IE, Settles ML. Vaginal microbiome
919 changes with levonorgestrel intrauterine system placement. Contraception. 2014
920 Aug;90(2):130–5.

921 76. Donders G, Bellen G, Janssens D, Van Bulck B, Hinoul P, Verguts J. Influence of
922 contraceptive choice on vaginal bacterial and fungal microflora. Eur J Clin Microbiol
923 Infect Dis Off Publ Eur Soc Clin Microbiol. 2017 Jan;36(1):43–8.

924 77. Fischer CN, Trautman EP, Crawford JM, Stabb EV, Handelsman J, Broderick NA.
925 Metabolite exchange between microbiome members produces compounds that
926 influence Drosophila behavior. eLife. 2017 09;6.

927 78. Bharwani A, Mian MF, Surette MG, Bienenstock J, Forsythe P. Oral treatment with
928 Lactobacillus rhamnosus attenuates behavioural deficits and immune changes in
929 chronic social stress. BMC Med. 2017 Jan 11;15(1):7.

930 79. Yang NJ, Chiu IM. Bacterial Signaling to the Nervous System through Toxins and
931 Metabolites. J Mol Biol. 2017 Jan 6;

932 80. Petrova MI, Reid G, Vaneechoutte M, Lebeer S. Lactobacillus iners: Friend or Foe?
933 Trends Microbiol. 2016 Nov 30;

934 81. Sweeney EL, Kallapur SG, Gisslen T, Lambers DS, Chougnet CA, Stephenson S-A, et al.
935 Placental Infection With Ureaplasma species Is Associated With Histologic
936 Chorioamnionitis and Adverse Outcomes in Moderately Preterm and Late-Preterm
937 Infants. J Infect Dis. 2016 Apr 15;213(8):1340–7.

938 82. Haggerty CL, Totten PA, Tang G, Astete SG, Ferris MJ, Norori J, et al. Identification of
939 novel microbes associated with pelvic inflammatory disease and infertility. Sex
940 Transm Infect. 2016 Sep;92(6):441–6.

941 83. MacKay AP, Kieki BA, Koonin LM, Beattie K. Tubal Sterilization in the United States,
942 1994-1996. Fam Plann Perspect. 2001;33(4):161–5.

943 84. Gondo F, da Silva MG, Polettini J, Tristao A da R, Peracoli JC, Witkin SS, et al. Vaginal
944 flora alterations and clinical symptoms in low-risk pregnant women. Gynecol Obstet
945 Invest. 2011;71(3):158–62.

946 85. Tierney L, Rossini A., Li N, Sevcikova H. Simple Network of Workstations. 2016.

947 86. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a
948 software environment for integrated models of biomolecular interaction networks.
949 Genome Res. 2003 Nov;13(11):2498–504.

950 87. Nayfach S, Pollard KS. Toward Accurate and Quantitative Comparative Metagenomics.
951 Cell. 2016 Aug 25;166(5):1103–16.

952 88. Bernstein MN, Doan A, Dewey CN. MetaSRA: normalized sample-specific metadata for
953 the Sequence Read Archive. bioRxiv. 2016 Nov 30;090506.

954 89. Bhadra A, Rao A, Baladandayuthapani V. Inferring network structure in non-normal
955 and mixed discrete-continuous genomic data. ArXiv160400376 Stat [Internet]. 2016
956 Apr 1 [cited 2017 Sep 29]; Available from: http://arxiv.org/abs/1604.00376

957

958 # Supporting information

959 **S1 File. Unfiltered, non-normalized, non-dichotomized microbiome counts and survey**

960 **responses (upon request)**

961 **S2 File. Survey questions and abbreviations**

962 **S3 File. Filtered, discretized microbiome counts and survey responses for all 60 variables**

963 **used in BN construction and analysis.**

964 **S4 File.** ***A priori*** **classifications, abbreviations, and BN analysis results (degree, closeness**

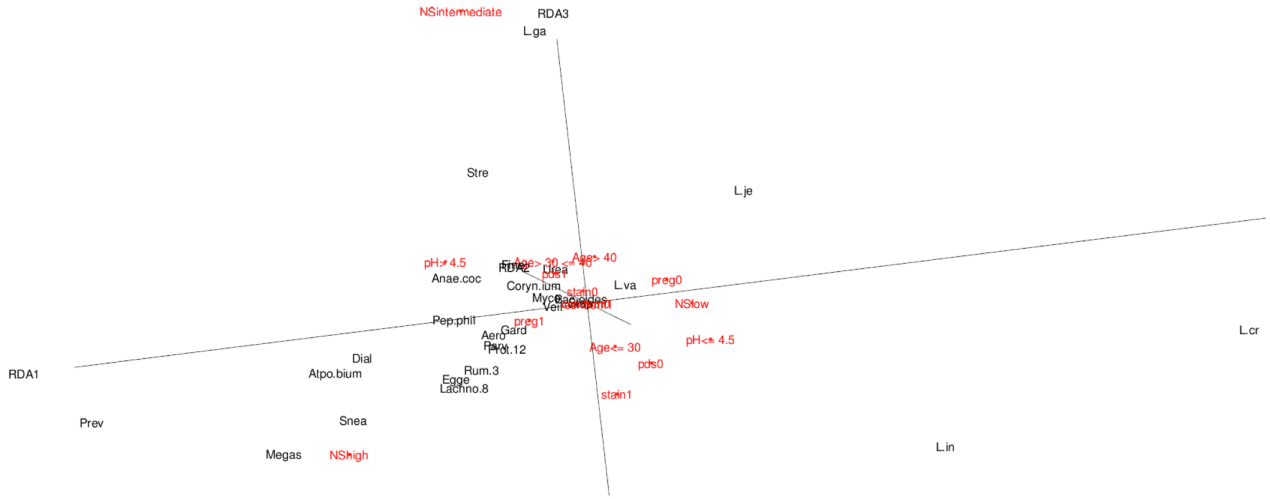
965 **centrality and betweenness centrality) for all 60 variables used in BN construction and**

966 **analysis.**

967 **S5 File. NMDS-ANOSIM testing results.**

968 **S6 File. NMDS plots of the six variables found to significantly influence the vaginal**

969 **microbiome using the ANOSIM test.**



970 **S7 Figure. 3D ordination plot based on PERMANOVA results.**

971 **S8 File. Figures describing the distribution of the abundances of taxa used in BN analysis.**

972 **Table 1. PERMANOVA permutation test results after stepwise model selection using the**

973 **AIC. PERMANOVA was used to assess the significance of associations between metadata**

974 **and the vaginal microbiome composition. P-values are based on 1000 permutations.**

Variable	Degrees of freedom	p-values
Nugent score	**2**	**0.001**
Previous pregnancy	**1**	**0.001**
Vaginal pH	**1**	**0.001**
Participant age	**2**	**0.012**
Use of menstrual pads during menstruation	**1**	**0.023**
Staining of underwear in previous 60 days	**1**	**0.027**
Use of male condoms as current birth control method	**1**	**0.066**

975