In silico trio-biomarkers for bacterial vaginosis revealed by species dominance network analysis

Zhanshan Sam Ma a,b,*, Aaron M. Ellison c

a Computational Biology and Medical Ecology Lab, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, China
b CAS Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China
c Harvard University, Harvard Forest, 324 North Main Street, Petersham, MA 01366, USA

ABSTRACT

BV (bacterial vaginosis) influences 20%–40% of women but its etiology is still poorly understood. An open question about the BV is which of the hundreds of bacteria found in the human vaginal microbiome (HVM) are the major force driving the vaginal microbiota dysbiosis. Here, we recast the question of microbial causality of BV by asking if there are any prevalent ‘signatures’ (network motifs) in the vaginal microbiome networks associated with it? We apply a new framework (species dominance network analysis by Ma & Ellison [2019]: Ecological Monographs) to detect critical structures in HVM networks associated with BV risks and etiology. We reanalyzed the 16 s-rRNA gene sequencing datasets of a mixed-cohort of 25 BV patients and healthy women. We failed to find any of these 15 trio-motifs in three additional cohorts of 1535 healthy women. Most member-species of the 15 trio motifs are BV-associated anaerobic bacteria (BVAB), Ravel’s community-state type indicators, or the most dominant species; virtually all species interactions in these trios are high-salience skeletons, suggesting that those trios are strongly connected ‘cults’ associated with the occurrence of BV. The presence of the trio motifs unique to BV may act as indicators for its personalized diagnosis and could help elucidate a more mechanistic interpretation of its risks and etiology. We caution that scarcity of large longitudinal datasets of HVM also limited further verifications of our findings, and these findings require further clinical tests to launch their applications.

ARTICLE INFO

Article history:
Received 19 January 2021
Received in revised form 5 May 2021
Accepted 9 May 2021
Available online 11 May 2021

Keywords:
BV (Bacterial vaginosis), BV-associated anaerobic bacteria (BVAB), Diversity-stability relationship (DSR), High-salience skeleton networks (HSN), Species dominance, Species dominance network (SDN)

1. Introduction

BV is a high-recurrence vaginal disease that occurs in 20%–40% of sampled women [54,30] depending on cohorts surveyed. It is associated with genital tract infections and many pregnancy complications [3,50,59,76], including pelvic inflammatory disease [6,72], premature rupture of membranes [12], intrauterine growth restriction, intrauterine fetal demise [9], chorioamnionitis [21], endometritis, preterm labor and delivery [29,33,71,74], postpartum infection, ectopic pregnancy [72], and tubal factor infertility. BV is also an independent risk factor for the acquisition and transmission of STDs (sexually-transmitted diseases) and HIV [36,64,73]. Beginning in the late 1990s, clinical microbiologists using culture-dependent technology (that could only cultivate and identify very limited number of bacteria) had already used ecological interpretations (particularly community diversity-stability relationships and species dominance) to interpret BV etiology [66]. The advent of NGS-based metagenomics has made it possible to detect large number of uncultivable microbes in the human vaginal microbiome (HVM) and greatly extended our capacity to holistically decipher the ecological processes underlying BV risks and etiology [7,17,23,36,50,51,53,59,75,18,52]. The recent characterization of BV as “a syndrome linked to various community types that cause somewhat similar physiological symptoms” by Ma et al. [36] reflects the limited ecological aspect of the state-of-the-art of research on BV etiology. This ecological syndrome perspective is only one of multiple characterizations of BV.

Abbreviations: ABV, asymptomatic bacterial vaginosis; BV, bacterial vaginosis; BVAB, BV-associated anaerobic bacteria; CPN, core/periphery network; CST, community state type; DSR, diversity-stability relationship; HEA, healthy treatment; HSN, high-salience skeleton network; HVM, human vaginal microbiome; MAO, most abundant species at OTU; MDO, most dominant species or OTU; OTU, operational taxonomic unit; SBV, symptomatic BV; SDN, species dominance network.

* Corresponding author at: CAS Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China.
E-mail address: ma@vandals.uidaho.edu (Z.S. Ma).

https://doi.org/10.1016/j.csbj.2021.05.020
2001-0370© 2021 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
as comprehensively reviewed by Rosca et al. [63] that includes disease, syndrome, infection, disorder, condition, etc. A reason we prefer the ecological characterization is because the first line of microbiome research is usually ecological studies.

The focus on a causal agent for BV asks: which of the hundreds of bacteria found in the HVM are the pathogens of BV? Even though Koch’s postulate of one pathogen to one disease has in some cases given way to a paradigm of polymicrobial disease etiology, studies of BV have neither identified a single causal pathogen nor a definite polymicrobial ‘signature’ (or what we call here a ‘cult’) of community composition associated with either a woman with a healthy vagina or one with BV. We were tempted to designate some cults of anaerobes in species dominance networks (SDNs) of the HVM as polymicrobial pathogens, but we detected the same cults in the SDN of healthy peers in the same investigated cohort (see “Results and Discussion” section) [19,26,36,37,41,42,34,59,76,62,63,53,61,38].

A series of longitudinal studies [19,59,61] together with other global efforts during the last decade [17,2,22,69,47,70,35,76,1,25,28,5,62,53,68,14,7,34,52,18,32] have greatly expanded our knowledge on HVM and BV. It had been argued that a lack of ‘good’ dominant species in the HVM, such as Lactobacillus spp., was associated with the increased risk of BV [27,53,55,61,8]. Another characteristic of the BV has been suggested to be associated with mixed multispecies communities with increased diversity and anaerobes [53,61]. This argument was extended to associate HVMs with high community “evenness” (no singularly dominant species) with elevated risk of BV. However, L. iners may still be abundant in BV patients and more even HVMs have been reported from asymptomatic women [36]. Ma & Ellison [41–42] re-examined the classic dominance (diversity)-stability relationship (DSR) paradigm [10,13,43–46], distinguished between stability and resilience, and introduced a new SDN-based framework for investigating the DSR in the healthy HVM.

Studies have found that recurrent BV might be associated with the presence and persistence of some bacteria, including Gardnerella spp., Atopobium vaginae, Mobiluncus spp., Bacteroides spp., Prevotella spp., Clostridiales spp., and Leptotrichia/Sneathia spp., Megapthera phyloyte 2, BV-associated bacteria, and Mobiluncus critisii [2,51,67] [77–82]. For example, Bradshaw et al. [2] reported that A. vaginae and Gardnerella spp. were present in >80% patients with recurrent BV. Meltzer et al. [81] suggested that there is significant relationship between the persistence of Mobiluncus critisii and recurrent BV. It has been well recognized that a universal pathogen(s) or a single diagnostic taxon for BV is unlikely. Furthermore, a single, universal dysbiotic community state is unlikely to cover all BV states. It has also been hypothesized that BV corresponds to one or more non-equilibrium states in the complex ecosystem that is the HVM and that there may be multiple paths to one or more non-equilibrium states in the complex ecosystem that is recurrent BV. Meltzer et al. [81] sequenced samples from vaginal communities collected daily for ten weeks from 25 women who were diagnosed as symptomatic BV (SBV: n = 15 women), asymptomatic BV (ABV: n = 6), or healthy (HEA: n = 4). In total, Ravel et al. [61] sequenced 1,657 samples (median = 67 per woman) and obtained 8,757,681 high-quality sequenced reads of the V1-V3 hypervariable region of 16S-rRNA genes, with a median of 5,093 reads per sample. Five of the 25 women were Caucasians and the remaining 20 were African American.

We used three additional 16S-rRNA datasets of healthy women originally collected by Ravel et al. [59], Gajer et al. [19], and Doyle et al. [14], respectively, to test the specificity of the 15 special trio motifs. A total of 1535 healthy women were sampled in these three additional datasets. The dataset of “32-healthy cohort” [19] consisted of the longitudinal sampling of 32 healthy women and has virtually the same data structure as the previously described “25-mixed cohort”. For this dataset, 32 SDNs were constructed to verify the findings from the primary dataset (25-mixed cohort). The “400-healthy cohort” dataset [59] consisted of the cross-sectional samples of 397 healthy women from four ethnic groups (Asian, Black, Hispanic, and White). For this dataset, we built a SDN for each of the four ethnic groups as well as a fifth SDN of the combined groups. Finally, the dataset of “1107-healthy cohort” [14] consisted of the cross-sectional samples of 1107 healthy, postpartum women. For this dataset, we built a single SDN for all individuals in the cohort.

2.2. Species dominance network (SDN) analysis

The SDN analysis we used was detailed in Ma & Ye [39] and Ma & Ellison [41,42]. Briefly, by noting that dominance (in terms of numbers of individuals or total biomass) of a one or a few taxa in a multi-species assemblage results in an uneven distribution of all the taxa in the assemblage, measures of dominance can be equated to measures of unevenness, and by inversion, evenness of the assemblage. Whereas species diversity and species evenness have been applied routinely to multi-species assemblages or “communities”, we have shown that measures of dominance can be applied to both communities and individual species [41,42]. The dominance metric that we derived provides a unified mathematical approach to measuring overall community dominance (= diversity), dominance by individual species [41,42], and relationships between dominance and community-wide stability and resilience. Additional insights were gained by developing techniques to detect trio motifs, core/periphery networks, and high-salience skeleton networks, and an approach to model network structure phenomenologically.

2.2.1. Trio motifs

A trio motif is a local network structure consisting of interactions among three nodes (species). In the present study, we detect the trios primarily consisting of BV-associated anaerobic bacteria (BVAB), and the structure and composition of trio motifs can have
significant implications for the stability of the HVMC and BV etiology [39,41,42]. Unlike standard correlation analysis of networks, descriptions and analyses of trio motifs can take into account multiple properties of network “nodes” (taxa or operational taxonomic units [OTUs]) or “edges” (links between taxa or OTUs). These properties include node types—the most dominant OTU (MDO), the most abundant OTU (MAO), or the “hub” (most connections to other OTUs)—or edge types—positive or negative between OTUs. The MAO is selected based on the relative species abundance level, but the MDO is determined by the species dominance, which measures the relative contribution of a species to community dominance. Both species dominance and community dominance were previously defined and demonstrated in Ma & Ellison [41,42].

2.2.2. Core/periphery and high-salience skeleton networks
Core/periphery [11,42] and skeleton network analyses [24,65,43] detect global (network- or community-wide) topological structures with significant implications for community stability or resilience. Core/periphery analyses focus on nodes, identifying more stable and highly connected “core” OTUs and less stable and often loosely-connected “peripheral” OTUs. In contrast, skeleton analyses address the edges, including the interaction paths connecting OTUs that form the “backbone” or high-salience skeletons in a SDN. Because the sets of nodes and edges fully determine the topology of a network, our SDN analyses [39,41,42] are complete—they address both nodes and edges (Fig. 1).

2.2.3. Analyses
We used the algorithms and computational procedures for describing and analyzing SDNs [41–42] to analyze the HVM datasets originally described by Ravel et al. [61]. Specifically, we used Spearman’s rank correlation coefficient ($\rho$) and for testing its significance, set $\alpha = 0.001$ after FDR (false discovery rate) adjustment. A SDN was considered valid if absolute $|\rho| > 0.5$. In addition, we set $\alpha = 0.05$ for testing significance of data analyzed using a Kruskal-Wallis one-way analysis of variance (ANOVA).

2.3. Verification
We used three additional 16S-rRNA datasets of healthy women originally collected by Ravel et al. [59], Gajer et al. [19], and Doyle et al. [14], respectively, to test the specificity of the 15 special trio motifs (see the results section for how they were identified). A total of 1535 healthy women were sampled in these three additional datasets. The dataset of “32-healthy cohort” [19] consisted of the longitudinal sampling of 32 healthy women (13 whites, 16 blacks, 1 Hispanic and 2 others) and has virtually the same data structure as the previously described “25-mixed cohort”. For this dataset, 32 SDNs were constructed to verify the findings from the primary dataset (25-mixed cohort). The “400-healthy cohort” dataset [59] consisted of the cross-sectional samples of 397 healthy women from four ethnic groups (Asian, Black, Hispanic, and White). For this dataset, we built a SDN for each of the four ethnic groups as well as a fifth SDN of the combined groups. Finally, the dataset of “1107-healthy cohort” [14] consisted of the cross-sectional sam-
amples of 1107 healthy, postpartum women in Malawi. For this data-
set, we built a single SDN for all individuals in the cohort.

3. Results

3.1. Basic properties of the SDNs

The MDO, MAO, and hubs of each subject are given in Table S1-1. In the whole cohort, the MDO of 11 (out of 25 or 44%) individuals was Lactobacillus iners. In the ABV group, the MDO of 5 (out of 6 or 83%) individuals was L. iners. The MDOs of the four HEA individuals were all different: L. iners, L. crispatus, L. jensenii and Bifidobacterium bifidum. The MAO of 11 (44%) subjects also was L. iners, and the MAO and MDO were the same OTU in 23 of the 25 individuals (all but #s112 and #s17). The hub OTUs differed among individuals and the various groups. Notably, a recent study proposed that there are at least 13 ‘genomic’ species within the currently taxonomically identified species G. vaginalis [75]. Some scholars have still retained the usage of the species name ‘G. vagi-
nalis’ (e.g., [20,57], while others have used Gardnerella spp. instead (e.g., [63]).

Network graphs for all 25 subjects are displayed in Fig. S1. Here, we illustrate network graphs for three exemplar subjects (Figs. 1–3) to illustrate basic network topology of the HVM. In these figures, we distinguish the MDO, MAO, and hub; one OTU may play two or all three of these roles. We also distinguish core (in cyan) and periphery (in azure blue) nodes; negative interactions (red) and positive ones (green). Thick edges identify the high-salience skeletons (salience-value \( S > 0.25 \), which measures the strengthen of species interactions).

Individual #s112 had different OTUs for its hub (Peptoniphilus), MDO (G. vaginalis), and MAO (L. iners) (Fig. 2). Although MAO of this individual was L. iners, she also had BV [61], in contrast with the current interpretation of BV etiology that suggests that high abundance of L. iners should lower susceptibility to BV [36]. However, the MDO of individual #s112 was G. vaginalis, an anaerobic species indicative of an altered vaginal environment. Some even consider G. vaginalis as the predominant cause of BV [4], including asymptomatic BV (ABV) [16,56]. The hub OTU of individual #s112 was Peptoniphilus, considered an indicator of higher risk of BV and of BV state type IV-B [19]. This example illustrates the utility of distinguishing among different roles (MDO, MAO, hub) in interacting networks of OTUs in the HVM [41,42].

Individual #s23 had different dominant OTUs: the hub was Prevotella buccalis and G. vaginalis was both the MAO and MDO (Fig. 3). This individual was diagnosed with ABV. Prevotella also is one of the indicator species of the BV state type IV-B [19]. Increased abundance of Prevotella in the HVM has been associated to BV, and Randis & Ratner [58] asserted that Gardnerella and Prevotella are “co-conspirators in the pathogenesis of bacterial vaginosis”.

Individual #s52, who was identified as healthy (HEA), had Lactobacillus jensenii as the MAO and MDO, and Clostridiales Family XI
Among the four sampled HEA women, the MDOs of three were *L. jensenii*. All negative interactions in the HVM network of #s52 were linked to *L. jensenii* and *L. gasseri*, which suggests their suppressive effects on BVABs and other opportunistic pathogens. Both *Lactobacillus* species are normal inhabitants of the lower reproductive tract in healthy women, whereas *Clostridiales* Family XI (incertae sedis) can be an opportunistic pathogen found in women diagnosed with BV [31].

The basic network properties of the individual SDNs of the 25-mixed cohort and the associated P-value from a Kruskal-Wallis one-way ANOVA are given in Table S1-2. Five network properties, average degree, diameter, average path length, network density and modularity, were significantly different between ABV and SBV individuals (*P* < 0.05). All other network properties were not significantly different among various groups (*P* > 0.05).

Across individuals, SDNs are variable, heterogeneous, and specific to individual women (Figs. 2-4, Fig S1, Table S1-2). As each woman has a unique HVM, HVM-associated BV requires personalized diagnosis and treatment. Individual SDN networks can reveal some unique aspects of individual subjects, but from standard network properties (Table S1-2), we can derive limited insights on either HVM heterogeneity or BV etiology.

### 3.2. Trio motifs and indicators of BV risk

We searched for the trio motifs in the individual SDNs by considering (i) both positive and negative interactions; (ii) the MDO, MAO, and BVAB; and (iii) two types of trio motifs: 3-node “Trios without handle” (the three-node trios, see Table S2-1A) and 4-node “Trios with MDO handle” (in which the MDO is connected to a three-node trio with one, two, or three links, see Table S2-1B). The total number of double-linked MDO (DLM) trios (in which the MDO is connected to three nodes with two links) was significantly different only between ABV and SBV individuals (*P* < 0.05). There were no significant differences in all other types of trios in any other comparisons (*P* > 0.05). We draw the following findings from this trio motif analysis:

(i) 951 trios occurred simultaneously in four or more women in the 25-mixed cohort. The reason we chose counting the trios that occurred in four or more subjects was that the minimal group size of BV, ABV and HEA groups in the cohort was 4 for the HEA group.
Among these 951 trios, one trio occurred in 15 women, another trio occurred in 12 women, four trios occurred in 10 women, six occurred in nine women, 16 occurred in eight women, 43 occurred in seven women, 114 occurred in six women, 255 occurred in five women, and 511 occurred in four women (see Fig. S2, Tables S2-2, S2-3 for more detailed information on the distribution of these 951 trios among ABV, BV and HEA groups).

As shown in Fig S2, 32% (301) of these 951 trios occurred exclusively in the SBV group, in which 12 trios occurred in near half of the subjects in the SBV group (Table 1). However, no trios were detected exclusively in ABV or HEA without also being detected in another group(s). Another 232 (24%) of the trios were detected in both SBV and ABV groups, in which three trios occurred in ten (50%) of the subjects in the BV (SBV + ABV) groups (Table 1). Finally, 340 trios (36%) occurred in all three groups.

Note that when we state a trio occurred in a group (e.g., SBV), it means that this trio occurred in at least one of the subjects in the group (SBV).

(ii) The 12 trios that occurred only in women with SBV potentially could be used as indicators of risk for BV (see Tables S2-3, S2-4 for their biomedical implications). Among these 12 trios, all three OTUs (nodes of the trio) in ten trios were BVABs and most were core species (see below for further explanation of the core species) of the individual SDNs. For example, trio #17 occurred in eight women with SBV subjects; its three nodes were Atopobium vaginae, BVAB2 and Parvimonas micra. Interactions between three BVABs were positive (cooperative), but the trio itself was inhibited by the MDO in seven of the eight women. Previous studies also have identified Atopobium and Parvimonas as indicators of BV-state CST (community state type) 4-B or CST-4B [19,40]. The two others of these 12 trios included L. iners and two BVABs.

(iii) Three trios were BV-only (BV = SBV + ABV) trios, occurring in 10 of the individuals diagnosed with either SBV or ABV (see Table S2-3 for their biomedical implications). Interactions among the three OTUs in these three trios were all positive. Trio #1—consisting of Dialister sp. type.2, Eggertella, and Veillonellaceae—occurred in 11 women with SBV subjects and 4 women with ABV. These three species also were the core species in the SDNs of all four women with ABV and 8 of the 11 with SBV. Trio #3—Anaero- coccus, Anaerococcus tetradius, and Peptoniphilus—occurred in 6 women with SBV and 4 with ABV. These three species were the core species in 7 of the 10 SDNs. Studies have identified Anaerococcus and Peptoniphilus as two other indicator species of the BV-state
Table 1

The 15 special trios, including the 12 trios exclusively occurred in the SBV group (SBV-only) and the 3 trios occurred in the BV-only (BV = SBV + ABV) (see Table S2-4 for biomedical information on the 15 trios), as well as the relevant core/periphery/skeleton information.

| No# | Figures of Trios | Trio members | BV Status | Subject ID | Average Salience Value** |
|-----|-----------------|--------------|-----------|------------|--------------------------|
| 17* | Atopobium vaginae BVAB2 Parvimonas micra | SBV | s128, s130, s135, s17, s40, s5, s53, s59 | 0.130 |
| 22  | Bifidobacteriaceae Parvimonas micra Peptoniphilus | SBV | s128, s130, s135, s17, s40, s53, s59 | 0.114 |
| 34* | Anaerococcus tetradus BVAB2 Peptoniphilus | SBV | s112, s116, s128, s135, s40, s5, s53 | 0.081 |
| 35  | Anaerococcus tetradus Gemella Peptoniphilus | SBV | s127, s128, s135, s17, s40, s5, s53 | 0.078 |
| 39* | Atopobium vaginae BVAB2 Eggerthella | SBV | s128, s130, s135, s40, s5, s53, s59 | 0.092 |
| 40* | Atopobium vaginae BVAB2 Peptoniphilus | SBV | s112, s116, s128, s130, s17, s40, s53 | 0.086 |
| 47* | BVAB2 Dialister sp. type 2 Eggerthella | SBV | s128, s130, s135, s40, s5, s53, s59 | 0.073 |
| 48* | BVAB2 Dialister sp. type 2 Peptoniphilus | SBV | s128, s130, s135, s40, s5, s53, s59 | 0.059 |
| 50  | BVAB2 Lactobacillus iners Peptoniphilus | SBV | s112, s130, s135, s15, s17, s53, s59 | 0.084 |
| 58* | Dialister sp. type 2 Parvimonas micra Peptoniphilus | SBV | s128, s130, s135, s35, s40, s53, s59 | 0.066 |

(continued on next page)
CST4-B [19]. Trio #5—Atopobium vaginae, Bifidobacteriaceae and Gardnerella vaginalis—was detected in 7 women with SBV and 3 with ABV. Among their 10 SDNs, Atopobium vaginae, Bifidobacteriaceae and Gardnerella vaginalis were detected, respectively, as the core species in 7, 8, and 5 of the SDNs. These three BV-only trios merit additional investigation as promising indicators of BV risk (Table S2-4). However, we consider they are more ambiguous indicators than the previous 12 SBV-only trios because they occurred in both SBV and ABV.

### Table 1 (continued)

| No# | Figures of Trios | Trio members | BV Status | Subject ID | Average Salience Value** |
|-----|-----------------|--------------|-----------|------------|--------------------------|
| 59* | ![DIagram](https://via.placeholder.com/150) | *Dialister sp. type 2 Parvimonas micra* *Prevotella genogroup 1* | SBV | s112, s128, s135, s35, s40, s53, s59 | 0.056 |
| 65 | ![Diagram](https://via.placeholder.com/150) | *Lactobacillus iners Parvimonas micra Peptoniphilus* | SBV | s130, s135, s15, s17, s35, s53, s59 | 0.083 |

The 3 special trios occurred in BV-only (BV = SBV + ABV) in 50% or more BV subjects

| 1 | ![Diagram](https://via.placeholder.com/150) | *Dialister sp. type 2* *Eggerthella Veillonellaceae* | SBV | s116, s127, s128, s130, s135, s15, s17, s35, s40, s5, s59 | 0.162 |
| 3* | ![Diagram](https://via.placeholder.com/150) | *Anaerococcus Anaerococcus tetradius Peptoniphilus* | SBV | s112, s127, s128, s130, s135, s15, s5, s27, s60, s77, s72 | 0.121 |
| 5 | ![Diagram](https://via.placeholder.com/150) | *Atopobium vaginae Bifidobacteriaceae Gardnerella vaginalis* | SBV | s127, s128, s130, s3, s35, s40, s59 | 0.275 |

The OTU frequencies of the core/periphery nodes***

| OTU ID | Core Frequency | Periphery Frequency | CST (community state type) | Is BVAB? |
|--------|----------------|---------------------|---------------------------|----------|
| Dialister sp. type 2 | 19 | 5 | CST4-B | BVAB |
| Peptoniphilus | 17 | 7 | CST4-B | BVAB |
| Veillonellaceae | 17 | 6 | CST4-B | BVAB |
| Bifidobacteriaceae | 14 | 8 | CST4-B | BVAB |
| Eggerthella | 14 | 6 | CST4-B | BVAB |
| Atopobium vaginae | 13 | 10 | CST4-B | BVAB |
| Anaerococcus | 13 | 9 | CST4-B | BVAB |
| BVAB2 | 11 | 12 | CST4-B | BVAB |
| Anaerococcus tetradius | 11 | 8 | CST4-B | BVAB |
| Parvimonas micra | 10 | 11 | CST4-B | BVAB |
| Gemella | 10 | 9 | CST4-B | BVAB |
| Lactobacillus iners | 10 | 8 | CST1 ~ CST3 | BVAB |
| Gardnerella vaginalis | 9 | 11 | CST4-B | BVAB |
| Prevotella genogroup 1 | 9 | 7 | CST4-B | BVAB |
| Mean | 12.642 | 8.357 | | |
| Standard Error | 3.225 | 2.098 | | |
| χ² test | | | χ² = 13.141 | P = 0.437 |

(i)** These 9 trios are all-BVAB (BV associated anaerobic bacteria) trios, i.e., all their three nodes are BVABs.

(ii)** It appears that these OTUs occurred more frequently in core than in the periphery (approximately 13 vs. 8), but the χ²-test indicates that the difference is not statistically significant.

(iii)** In the column of figures of trios, pink nodes in the trios are BVABs, and green nodes are non-BVABs.

(iv)** The average salience values at the whole network level (see Table S4-4) for the test results.

The average skeleton salience (interaction strength) values of these special trios are significantly higher than the other 15 trios consisting of 14 OTUs, 10 out of which are BVABs, and 8 out of which are CST indicators.
Zhanshan Sam Ma and A.M. Ellison
Computational and Structural Biotechnology Journal 19 (2021) 2979–2989

is no single trio that was detected in all-individual SDNs of the 25-mixed cohort, which simply displays enormous inter-subject heterogeneity of the HVM. For this, here we choose to detect trios occurred in simple majority (≥50%) of each diagnostic group (SVB, ABV or HEA).

In the SBV group, there were 22 trios detected in > 50% of the women (≥7 out of 15), in which only one trio (trio #1) occurred in 73% (11/15) women, and 3 occurred in 53% (8/15) women. Among these 22 trios detected in the SBV group: 12 trios occurred exclusively in the SBV group, as described in (ii) previously; nine trios occurred in both the SBV and ABV groups, including trios #1, #3 and #5 described in (iii) previously; and one trio (trio #6) occurred in all three diagnostic groups including HEA group.

In the ABV group, there were 21 trios detected in 50% or more subjects (≥3 out of 6), in which one trio (#2) occurred in 80% (5/6) subjects, three trios (#1, #3 and #4) occurred in 70% (4/6) subjects. Among these 21 trios, 10 trios also occurred in both the SBV and ABV groups, including #1, #3 and #5 discussed in (iii), and 11 trios occurred in all three diagnostic groups.

In the HEA group, there were 244 trios detected in ≥50% of the subjects. However, only 2 trios (#4 & #72) out of the 244 were detected in 75% (3/4) subjects. Among these 244 trios, 201 trios including (#4 & #72) were also detected in the SBV and ABV groups, and the remaining 43 occurred in all three diagnostic groups.

In summary, we identified 15 special trios, including 12 SBV-only trios occurred exclusively in the SBV group and 3 BV-only trios occurred in ≥50% of the BV subjects only (BV = ABV + SBV groups). Table S2-4 provides brief biomedical information on the members of the 15 trios. These trios are of potentially significant importance for assessing BV risks and for investigating BV etiology.

3.3. Core/periphery species in the SDNs

Properties of the core/periphery structures for the individual SDNs of the 25-mixed cohort are given in Table S3-1. There were no significant differences in the network properties between women with BV (SBV + ABV) and those without (HEA). The core strength (ρ), its size as the proportion of connected network nodes (C/[C + P]), and degree of nestedness (S) were significantly larger in SBV than ABV individuals (P < 0.05).

Regression analysis of overall network dominance as a function of core/periphery dominance revealed the following relationship:

\[
\ln(\text{Network}) = 1.97 + 0.05\ln(\text{Core}) + 0.48\ln(\text{Periphery}) \quad (r = 0.85, P \ll 0.001)
\]  

(1)

The above model clearly indicates that periphery has a larger effect on community-wide dominance given that the scaling parameter of periphery is nearly 10 times that of core (0.48 vs. 0.05). Back-transforming these models gives the exponential relationship between network dominance and core/periphery dominance:

\[
\text{Network} = 7.17(\text{Periphery})^{0.48}(\text{Core})^{0.05}
\]  

(2)

The core/periphery status of individual OTUs varied among individuals (Tables S3-2A, S3-2B). Examination of frequency distributions of OTUs of the HVM in the 25-mixed cohort revealed that although many species occurred in the core (104 OTUs) or periphery (146 OTUs), only a small number of them were exclusive to either (3 exclusive to the core, 42 exclusive to the periphery). Note that the core/periphery status is not fixed in the cohort and can be individual-specific. Interestingly, one of the most commonly observed and also the most well-known species, \textit{L. iners}, in the HVM is only the 16th most frequent core species and it occurred in 10 individuals (SDNs) only. It also occurred in the periphery of 8 SDNs (8 subjects) and ranked after the 50th. This indicates that \textit{L. iners} may not be a universal ‘power’ player in the HVM.

Distribution in the core and periphery of BVAB differed significantly (P < 0.0001, χ² test; Table S3-3) and were more frequent in the periphery. Table S5 lists all the BVABs recovered from the samples (see also [61]). Table S3-3 also shows that although the BVABs may occur in either core or periphery, they occurred more frequently in the periphery, 314 times in cores vs. 429 in peripheries, or 37% more in the periphery. This finding echoes the previous finding that the periphery is more influential than the core to the network dominance as a whole.

Finally, we analyzed shared core/periphery networks by comparing pairs of subjects from two different groups (ABV, SBV, or HEA; Table S3-4). Using two different algorithms [43]—the A1 algorithm that reshuffles reads and the A2 algorithm that reshuffles samples—we found that the observed numbers of shared core (or periphery) species were significantly lower than those expected by chance among all three groups (P < 0.001 all cases). That is, there exist group-specific, unique core (or periphery) species for each of the SBV, ABV and HEA groups (Table S3-5).

3.4. High-salience skeletons in the SDNs

The salience of skeletons measures the strength of interactions in terms of species co-dominance. Table S4-1 gives the properties of the high-salience skeletons in all the SDNs in the 25-mixed cohort; those with salience value ≥ 0.5 are reported in Table S4-2. In the latter, the two nodes (OTUs) connected by the skeleton are MDOs, MAOs, hubs, BV-state indicators, or BVABs. We also tested the number of shared skeletons between each pair of subjects from the three diagnostic groups (Table S4-3). The average of shared skeletons between the HEA and SBV group, between the HEA and ABV group, between the ABV and SBV group, was 79, 91, and 129, respectively (Table S4-3B), with relative similarity ordered as ABV vs. SBV > ABV vs. HEA > SBV vs. HEA. This suggests that BV can change the strength of network links relative to those observed in healthy individuals.

4. Conclusions and discussion

Basic SDN analysis confirmed that human vaginal microbiomes are highly heterogeneous among individuals (see also [42]). It is unlikely that either a universal BV pathogen (diagnostic OTUs or their cults) or a universal dysbiotic BV state could be discovered due to the extreme heterogeneity. Nevertheless, our results suggest that it seems possible to discover a set of local network structures such as the trio motifs across a cohort (population) of BV patients. Core/periphery and high-salience skeleton network analyses found additional complexities and did not identify any universal clear-cut network structures or properties for diagnosing BV or interpreting BV etiology. Nonetheless, the 15 BV-only trio motifs (especially the 12 found exclusively in women with SBV) have promise for assessing and predicting BV risks and provide new insights about its mechanisms and etiology. In addition to using these special trio motifs as indicators of BV risk, individualized (personalized) analysis of each SDN can offer further mechanistic interpretations of the BV status of a particular individual (Figs. 2-4, Fig S1).

Although the presence of certain species (such as \textit{Lactobacillus} spp., \textit{G. vaginalis}), or their co-occurrence may not be sufficient evidence for determining BV status, but the existence of their interactions (special trio motifs) and interaction strengths (measured by skeleton salience) do provide supporting diagnostic evidence. 


Further insights can be shed on the ecological mechanisms ofBV by the distinguishing core and periphery OTUs, or by the identifications of critical pathway of species interactions (network backbones consisting of high-salience skeletons). Although the vaginal microbiomes of different individuals may have different levels of stability or resilience to various disturbances including BV, BV may indeed change the core strength/size and nestedness of SDNs and the strength of species interactions within an SDN, resulting in a change in its stability or resilience. Therefore, the core/periphery network and high-salience skeleton network [42] further complemented the findings and insights from special trio motif and inductive inspections on individual SDNs. In conclusion, the 15 BV-only trio motifs, their nodes and links, and the core/periphery/high-salience skeleton structures of their associated SDNs are sufficiently unique to act as indicators of community composition and signaling the potential of dysbiosis associated with BV. Although we are still far from revealing clear-cut risk or diagnostic indicators, our study presents important material candidates (15 BV-only trio motifs) and tools (trio-motif detection technique, augmented with core /periphery/skeleton network analyses) for further deepening our understanding on BV risks and etiology.

Two limitations exist in this study. First, scarcity of large longitudinal datasets of HVM limited further verifications of our findings, and our findings require further wet-lab biomedical and clinical tests to launch any possible applications. Second, the effects of ethnicity was not addressed, again due to the data limitation. Nonetheless, we have tried our best to use additional datasets to cross-verify our findings, as explained previously and summarized below. It should be reiterated that these verifications are imperfect, although they do offer supporting evidence to our findings.

Finally, we asked how specific are the 15 BV-only trio-motifs? We answered this question by testing whether the 15 special trio-motifs (listed in Table 1) could be detected in cohorts of healthy women. We used three additional 16S-rDNA datasets of 1535 healthy women to assess specificity. Although there were differences among the three healthy-cohort datasets (e.g., longitudinal vs. cross-sectional sampling designs), we were only interested in determining if we detected the special BV-only trio motifs in species dominance networks of healthy individuals. None of the 15 special trio-motifs were detected in any of the species dominance networks of the healthy women, regardless of sampling design (Table S5). Thus, it appears that the 15 trio-motifs could be a highly specific indicator of BV to be used for its personalized diagnosis and treatment.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study received funding from the following sources: A National Natural Science Foundation (NSFC) Grant (No. 31970116); Cloud-Ridge Industry Technology Leader Grant; A China-US International Cooperation Project on Genomics/Metagenomics Big Data. We appreciate Wendy Li and Lianwei Li of the Chinese Academy of Sciences for their computational support. The funding sources do not play any roles in interpreting the results or inferring the conclusions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.05.020.

References

[1] Aagaard K, Riehle K, Ma J, et al. A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. PLoS One 2012;7:e36466.
[2] Bradshaw CS, Tabrizi SN, Fairley CK, et al. The association of Atopobium vaginae and Gardnerella vaginalis with bacterial vaginosis and recurrence after oral metronidazole therapy. J Infect Dis 2006;194:828–36.
[3] Bradshaw CS, Brotman RM. Making inroads into improving treatment of bacterial vaginosis—striving for long-term cure. BMC Infect Dis 2015;15:292.
[4] Brotman RM, Rosenhein P, Pascal A, et al. High Atopobium vaginae and Gardnerella vaginalis vaginal loads are associated with preterm birth. Clin Infect Dis 2015;60(6):860–7.
[5] Brotman RM, Bradford LD, Conrad M, et al. Association between Trichomonas vaginalis and vaginal bacterial community composition across reproductive-age women. Sex Transm Dis 2013;39:807–12.
[6] Brunham RC, Gottlieb SL, Paavonen J. Pelvic inflammatory disease. N Engl J Med 2004;321(5):395–403.
[7] Castello EK, Stagaman K, Dethlefsen L, et al. The application of ecological theory toward an understanding of the human microbiome. Science 2012;336(6086):1255–62.
[8] Coemery P, London A, Wu LY, et al. Structure and dynamics of core/periphery networks. J Complex Networks 2013;1:93–123.
[9] Darwish A, Elshar EM, Hamadeh SM, et al. Treatment options for bacterial vaginosis in patients at high risk of preterm labor and premature rupture of membranes. J Obstet Gynaecol Res 2007;33(6):781–7.
[10] Donohue I, Hillebrand H, Montoya JM, et al. Navigating the complexity from an ecological perspective. Transl Res 2012;160:267–82.
[11] Doyle R, Gondwe A, Fan Y, et al. A Lactobacillus-Deficient Vaginal Microbiota Dominates Postpartum Women in Rural Malawi. Appl Environ Microbiol 2018;84(6):e02150–e2217.
[12] Durugo B, Nyengidiki TK, Bassey G, et al. Bacterial vaginosis among women with tubal factor infertility in Nigeria. Int J Gynaecol Obstet 2015;131(2):133–6.
[13] Eren AM, Zozaya M, Taylor CM, Dowd SE, Martin DH, Ferris MJ. Exploring the diversity of Gardnerella vaginalis in the genitourinary tract microbiota of mucous-coupled couples through subtle nucleotide variation. PLoS One 2011;6:e26732.
[14] Fredricks DN, Fiedler TL, Marrazzo JM. Molecular identification of bacteria associated with bacterial vaginosis. N Engl J Med 2005:353:1899–911.
[15] Fredricks DN. Molecular methods to describe the spectrum and dynamics of the vaginal microbiota. Anaerobe 2011;17(4):191–5.
[16] Gajer P, Brotman RM, Bail G, et al. Temporal dynamics of the human vaginal microbiome. Anaerobe 2011;17(4):191–5.
[17] García EM, Kranzaukienė V, Kohlinski JR, et al. Interaction of gardnerella vaginalis and vaginolysin with the apical versus basolateral face of a three-dimensional model of vaginal epithelium. Infect Immun 2019;87(4).
[18] Gibbs RS. Chloroammonitosis and bacterial vaginosis. Am J Obstet Gynecol 1993;169:460–2.
[19] Gibbs RS. Asymptomatic bacterial vaginosis: is it time to treat?. Am J Obstet Gynecol 2007;196:455–6.
[20] Gilbert NM, Lewis WC, Li G, et al. Gardnerella vaginalis and Prevotella bivia trigger distinct and overlapping phenotypes in a mouse model of bacterial vaginosis. J Infect Dis 2019;220(7):1099–108.
[21] Grady D, Thiemann C, Brockmann D. Robust classification of salient links in complex networks. Nat Commun 2012;3:854.
[22] Guendou FA, Van Damme L, Mirembe F, et al. Intermediate vaginal flora is associated with HIV prevalence as strongly as bacterial vaginosis in a cross-sectional study of participants screened for a randomized controlled trial. Sex Transm Infect 2012;88:545–51.
[23] Hickey RJ, Zhou X, Pierson JD, et al. Understanding vaginal microbe complexity from an ecological perspective. Transl Res 2012;160:267–82.
[24] Jakobsson T, Forsum U. Lactobacillus iners: a marker of changes in the vaginal flora?. J Clin Microbiol 2007;45:3145.
[25] Jespers V, Menten J, Smet H, et al. Quantification of bacterial species of the vaginal microbiome in different groups of women, using nucleic acid amplification tests. BMC Microbiol 2012:12:83.
[26] Klebanoff MA, Brotman RM. Treatment of bacterial vaginosis to prevent preterm birth. Lancet 2018;392(10160):2141–2.
[27] Koumans EH, Kendrick JS. Preventing adverse sequelae of bacterial vaginosis: a public health program and research agenda. Sex Transm Dis 2001;28(5):292–7.
