



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

Zhanshan Sam Ma<sup>a,b,\*</sup>, Aaron M. Ellison<sup>c</sup>

<sup>a</sup> Computational Biology and Medical Ecology Lab, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, China

<sup>b</sup> CAS Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China

<sup>c</sup> Harvard University, Harvard Forest, 324 North Main Street, Petersham, MA 01366, USA



## 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)  
Community dominance  
Core/periphery network (CPN)  
Diversity-stability relationship (DSR)  
High-salience skeleton networks (HSN)  
Species dominance  
Species dominance network (SDN)

## 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. In these datasets, we detected 15 trio-motifs that occurred exclusively in BV patients. 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.

© 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/>).

## 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 [3,49,53,60,76], 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], postpar-

*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 or 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](mailto: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/>).

tum infection, ectopic pregnancy [72], and tubal factor infertility [15]. 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 [7,17,23,36,50,51,53,59,76,18,52]. 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

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 [17,2,22,69,48,47,70,35,76,1,25,28,5,62,53,68,14,7,34,52,18,32] 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., *Megasphaera* phylotype 2, BV-associated bacteria, and *Mobiluncus cristicus* [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 cristicus* 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 types of BV states (henceforth "BV states") (e.g., [36,17,59]). We further postulate that despite their multiplicity, BV states are likely to be ecologically equivalent in terms of their decreased ecological stability or resilience [42,44–46]. With these hypotheses as a starting point, we conceive that it is feasible to detect certain 'signatures' of BV states. If the HVM is modeled as a complex network, finding a signature of BV states is reducible to finding a network motif (see Material and Methods, below for a brief presentation of motifs and other network characteristics; and [42] for a more detailed discussion of them). We follow the principle of parsimony and search for simple trio motifs (a basic local structure of bacterial interactions, consisting of three taxa) in complex networks [39]. We also harness other characteristics of complex ecological networks (including core/periphery and high-salience structures) point at underlying processes (mechanisms) driving dynamics of vaginal microbiomes and effects of BV as ecological disturbances

or perturbations [41–42]. Integrated together, our motif detection and SDN analyses may be developable as predictors of BV risks and its etiology. Fig. 1 illustrates the primary objectives and approaches of this study.

## 2. Material and Methods

### 2.1. The human vaginal microbiome (HVM) datasets

We used the "25-mixed cohort" (including both BV patients and healthy controls) dataset originally described by Ravel et al. [61] for primary species dominance network (SDN) analysis. Briefly, Ravel et al. [61] 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 ( $\approx$ 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

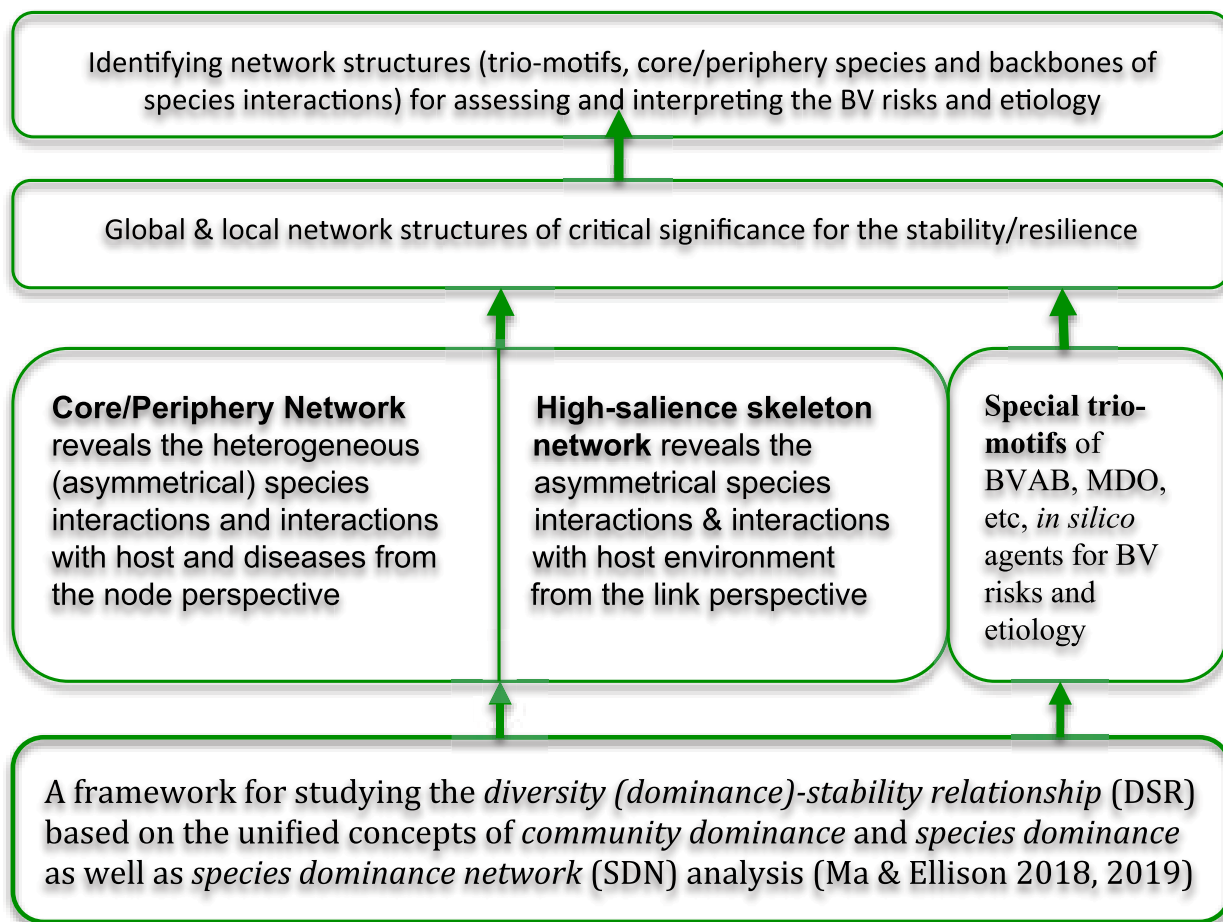


Fig. 1. A diagram showing the primary objectives and approaches to achieving the objectives.

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 accounts 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 focuses 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 ( $R$ ) and for testing its significance, set  $\alpha = 0.001$  after FDR (false discovery rate) adjustment. A SDN was considered valid if absolute  $|R| \geq 0.5$ . 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-

ples of 1107 healthy, postpartum women in Malawi. For this dataset, 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 *Gardnerella vaginalis*, but in the SBV group, the MDO of 9 (out of 15 or 60%) 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. vaginalis*” (e.g., [20,57], while others have used *Gardnerella* spp. instead (e.g., [63]).

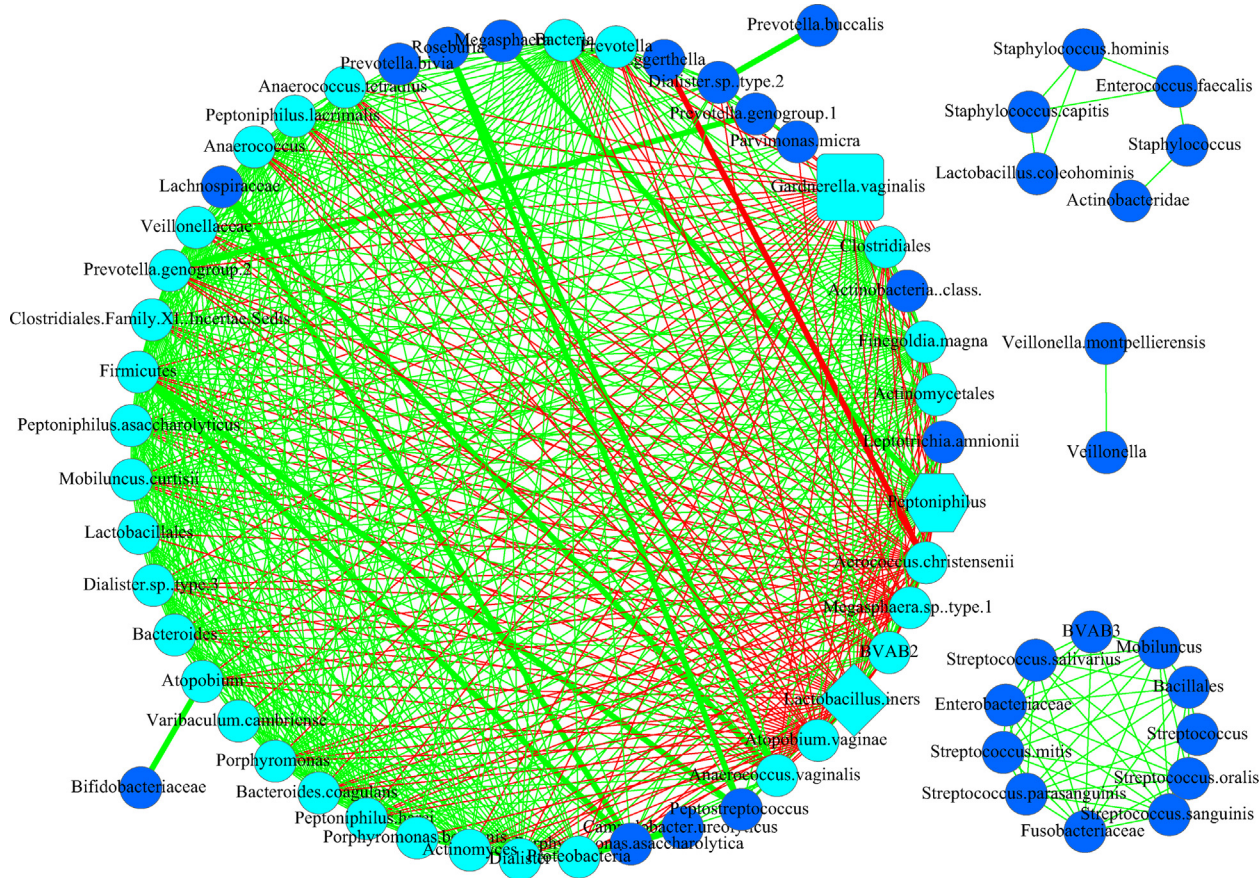
Network graphs for all 25 subjects are displayed in Fig. S1. Here, we illustrate network graphs for three exemplar subjects (Figs. 1-3) to illustrate basic network topology of the HVM. In these figures, we distinguish the MDO, MAO, and hub; one OTU may play two or

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  $S \geq 0.25$ , which measures the strengthen of species interactions).

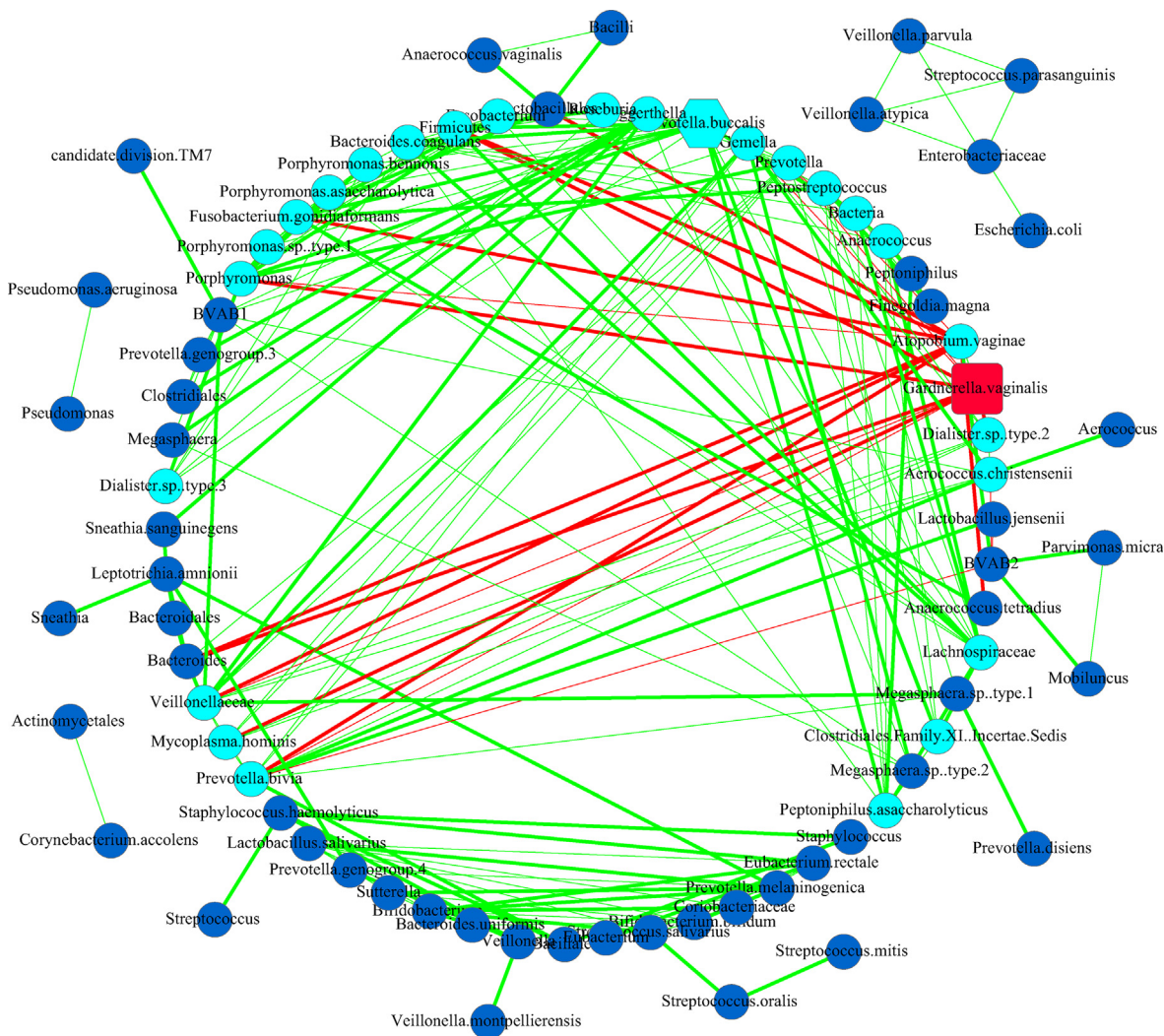
Individual #s112 had different OTUs for its hub (*Peptoniphilus*), MDO (*G. vaginalis*), and MAO (*L. iners*) (Fig. 2). Although MAO of this individual was *L. iners*, she also had BV [61], in contrast with the current interpretation of BV etiology that suggests that high abundance of *L. iners* should 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



**Fig. 2.** The SDN network graph for individual #s112 in the 25-mixed cohort (hub = *Peptoniphilus*, MAO = *Lactobacillus iners*; MDO = *Gardnerella vaginalis*; BV Status = SBV; hub, MAO and MDO all belong to the core; core nodes are in cyan color, and periphery nodes are in azure; rectangle in cyan represents that the node is both core and MDO; hexagon in cyan represents that the node is both core and hub; diamond in cyan represents that the node is both core and MAO; edges in green are positive correlations; edges in red are negative correlations; thicker edges are skeletons). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 3.** SDN network graph for individual #s23 in the 25-subjects cohort (hub, MDO & MAO are the same node, i.e., *Gardnerella vaginalis*, hub = *Prevotella buccalis*; MDO, MAO and hub all belong to the core; core nodes are in cyan color, and periphery nodes are in azure; BV-status = ABV; rectangle in red represents that the node plays triple role of core, MAO and MDO; hexagon in cyan represents that the node is both core and hub; edges in green are positive correlations; edges in red are negative correlations; thicker edges are skeletons). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(*incertae sedis*) as her hub species (Fig. 4). 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).

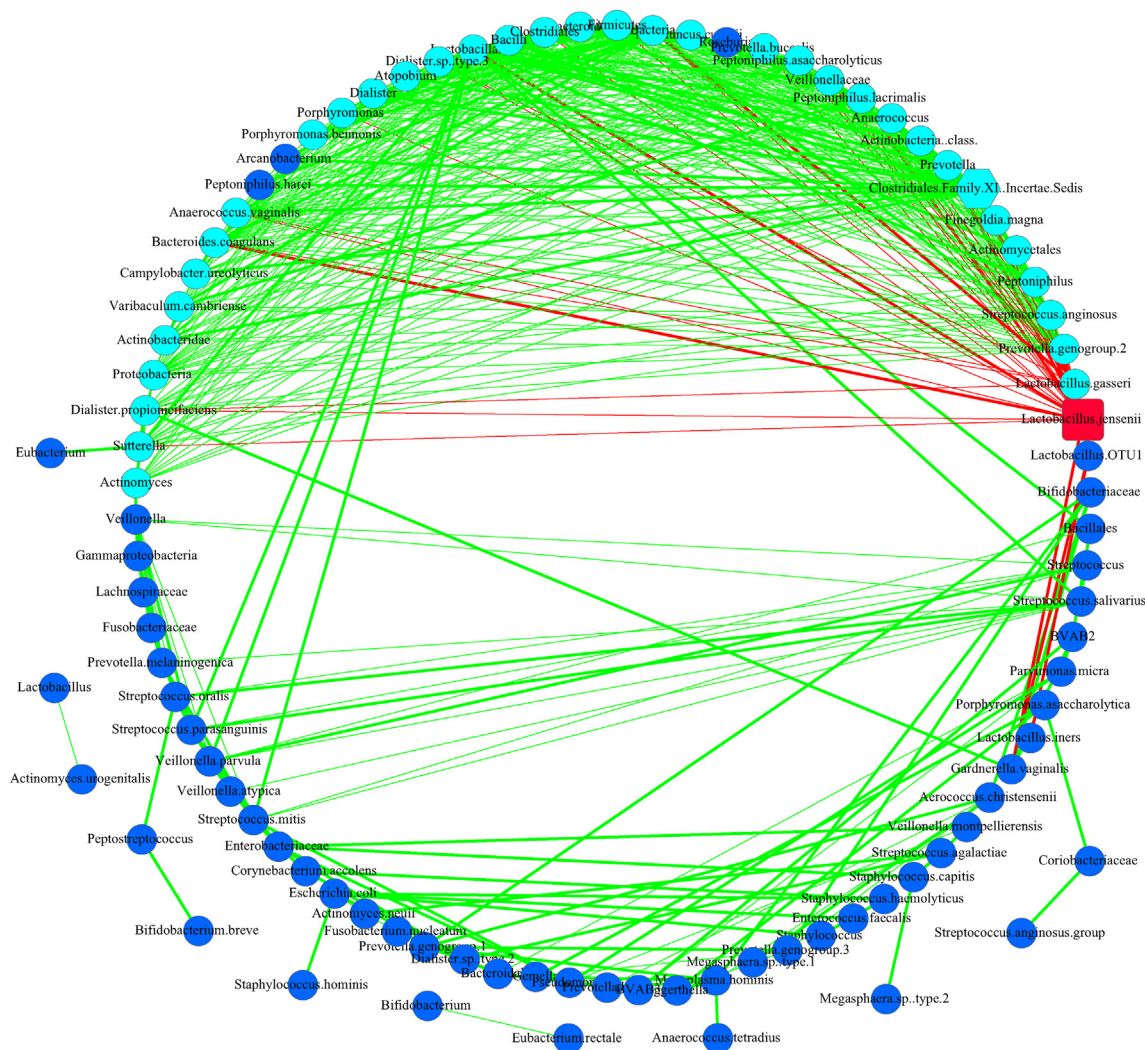
Across individuals, SDNs are variable, heterogeneous, and specific to individual women (Figs. 2-4, Fig S1, Table S1-2). As each woman has a unique HVM, HVM-associated BV requires personalized diagnosis and treatment. Individual SDN networks can reveal some unique aspects of individual subjects, but from standard net-

work 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 Tables 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 trio 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.



**Fig. 4.** SDN network graph for individual #s52 in the 25-subjects cohort (hub, MDO & MAO are the same node, i.e., *Lactobacillus jensenii*, hub = Clostridiales Family XI (incertae sedis); MDO, MAO and hub all belong to the core; core nodes are in cyan color, and periphery nodes are in azure; BV-status = HEA); rectangle in red represents that the node plays triple role of core, MAO and MDO; hexagon in cyan represents that the node is both core and hub; edges in green are positive correlations; edges in red are negative correlations; thicker edges are skeletons). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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% (30/91) 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. *L. iners* interacted negatively with the two BVABs in all the SDNs that contained these two trios, corroborating previous observations of inhibitory effects of *L. iners* on 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, *Eggerthella*, 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—*Anaerococcus*, *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.

Trios					
No#	Figures of Trios	Trio members	BV Status	Subject ID	Average Saliense Value**
The 12 special trios occurred exclusively in the SBV group (occurred in 50% or more SBV subjects)					
17*		<i>Atopobium vaginae</i> BVAB2 <i>Parvimonas micra</i>	SBV	s128, s130, s15, s17, s40, s5, s53, s59	0.130
22		<i>Bifidobacteriaceae</i> <i>Peptoniphilus</i> <i>Parvimonas micra</i>	SBV	s128, s130, s135, s15, s17, s40, s53, s59	0.114
34*		<i>Anaerococcus tetradius</i> BVAB2 <i>Peptoniphilus</i>	SBV	s112, s116, s128, s135, s40, s5, s53	0.081
35		<i>Anaerococcus tetradius</i> <i>Peptoniphilus</i> <i>Gemella</i>	SBV	s127, s128, s135, s17, s40, s5, s53	0.078
39*		<i>Atopobium vaginae</i> BVAB2 <i>Eggerthella</i>	SBV	s128, s130, s135, s40, s5, s53, s59	0.092
40*		<i>Atopobium vaginae</i> BVAB2 <i>Peptoniphilus</i>	SBV	s112, s116, s128, s130, s17, s40, s53	0.086
47*		Dialister sp. type 2 BVAB2 <i>Eggerthella</i>	SBV	s128, s130, s135, s40, s5, s53, s59	0.073
48*		Dialister sp. type 2 BVAB2 <i>Peptoniphilus</i>	SBV	s128, s130, s135, s40, s5, s53, s59	0.059
50		BVAB2 <i>Lactobacillus iners</i> <i>Peptoniphilus</i>	SBV	s112, s130, s135, s15, s17, s53, s59	0.084
58*		Dialister sp. type 2 <i>Parvimonas micra</i> <i>Peptoniphilus</i>	SBV	s128, s130, s135, s35, s40, s53, s59	0.066

(continued on next page)

Table 1 (continued)

Trios						
No#	Figures of Trios	Trio members	BV Status	Subject ID	Average Saliency Value**	
59*		<i>Dialister sp. type 2</i> <i>Parvimonas micra</i> <i>Prevotella genogroup 1</i>	SBV	s112, s128, s135, s35, s40, s53, s59	0.056	
65		<i>Lactobacillus iners</i> <i>Parvimonas micra</i> <i>Peptoniphilus</i>	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		<i>Dialister sp. type 2</i> <i>Eggerthella</i> <i>Veillonellaceae</i>	SBV ABV	s116, s127, s128, s130, s135, s15, s17, s35, s40, s5, s59 s27, s60, s77, s82	0.162	
3*		<i>Anaerococcus tetradius</i> <i>Peptoniphilus</i> <i>Anaerococcus</i>	SBV ABV	s112, s127, s128, s135, s15, s5 s12, s27, s60, s77	0.121	
5		<i>Atopobium vaginae</i> <i>Bifidobacteriaceae</i> <i>Gardnerella vaginalis</i>	SBV ABV	s127, s128, s130, s3, s35, s40, s59 s27, s60, s77	0.275	
The OTU frequencies of the core/periphery nodes***						
OTU ID	Core Frequency	Periphery Frequency	CST (community state type)	Is BVAB?		
<i>Dialister sp. type 2</i>	19	5	–	BVAB		
<i>Peptoniphilus</i>	17	7	CST4-B	BVAB		
<i>Veillonellaceae</i>	17	6	–	BVAB		
<i>Bifidobacteriaceae</i>	14	8	–	–		
<i>Eggerthella</i>	14	6	–	BVAB		
<i>Atopobium vaginae</i>	13	10	CST4-B	BVAB		
<i>Anaerococcus</i>	13	9	CST4-A	BVAB		
BVAB2	11	12	–	BVAB		
<i>Anaerococcus tetradius</i>	11	8	CST4-A	BVAB		
<i>Parvimonas micra</i>	10	11	CST4-B	BVAB		
<i>Gemella</i>	10	9	–	–		
<i>Lactobacillus iners</i>	10	8	CST1 ~ CST3	–		
<i>Gardnerella vaginalis</i>	9	11	CST4-B	BVAB		
<i>Prevotella genogroup 1</i>	9	7	CST4-B	BVAB		
Mean	12.642	8.357				
Standard Error	3.225	2.098				
$\chi^2$ test***	$\chi^2 = 13.141$	$P = 0.437$				

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

(iii) It appears that these OTUs occurred more frequently in core than in the periphery (approximately 13 vs. 8), but the  $\chi^2$ -test indicates that the difference is not statistically significant.

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

(ii) The average saliency values at the whole network level (see Table S4-4) for the test results.

\*\* The average skeleton saliency (interaction strength) values of these special trios are significantly higher than

\*\*\* The 15 trios consisted of 14 OTUs, 10 out of which are BVABs, and 8 out of which are CST indicators.

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.

(iv) We further identified the trios occurred in 50% or more subjects in each of the three diagnostic groups (see Table S2-5). Note that the term *occurrence* is used non-exclusively here. In fact, there



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 ( $\geq 50\%$ ) of each diagnostic group (SBV, ABV or HEA).

In the SBV group, there were 22 trios detected in  $> 50\%$  of the women ( $\geq 7$  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 ( $\geq 3$  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 (healthy) group, there were 244 trios detected in  $\geq 50\%$  ( $\geq 2$  out of 4) 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 both the HEA and SBV groups.

(v) 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  $\geq 50\%$  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 ( $\rho$ ), 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 ( $P \leq 0.05$ ).

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

$$\ln(\text{Network}) = 1.97 + 0.05\ln(\text{Core}) + 0.48\ln(\text{Periphery}) \quad (r = 0.85, P \ll 0.001) \quad (1)$$

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:

$$\text{Network} = 7.17(\text{Periphery})^{0.48}(\text{Core})^{0.05} \quad (2)$$

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, *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 *L. iners* may not be a universal ‘power’ player in the HVM.

Distribution in the core and periphery of BVAB differed significantly ( $P < 0.0001$ ,  $\chi^2$  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  $\geq 0.5$  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 for the BV status of a particular individual (Figs. 2-4, Fig S1).

Although the presence of certain species (such as *Lactobacillus* spp., *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 of BV 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-rRNA 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] Bretelle F, Rozenberg 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 LL, Conrad M, et al. Association between *Trichomonas vaginalis* and vaginal bacterial community composition among reproductive-age women. *Sex Transm Dis* 2012;39:807–12.
- [6] Brunham RC, Gottlieb SL, Paavonen J. Pelvic inflammatory disease. *N Engl J Med* 2015;372:2039–48.
- [7] Castro J, Machado D, Cerca N. Unveiling the role of *Gardnerella vaginalis* in polymicrobial Bacterial Vaginosis biofilms: the impact of other vaginal pathogens living as neighbors. *ISME J* 2019;13(5):1306–17.
- [8] Cohen CR, Wierzbicki MR, French AL, et al. Randomized trial of lactin-V to prevent recurrence of bacterial vaginosis. *N Engl J Med* 2020;382:1906–15.
- [9] Cottrell BH, Shannahan M. Maternal bacterial vaginosis and fetal/infant mortality in eight Florida counties, 1999 to 2000. *Public Health Nurs* 2004;21(5):395–403.
- [10] Costello 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.
- [11] Csermely P, London A, Wu LY, et al. Structure and dynamics of core/periphery networks. *J Complex Networks* 2013;1:93–123.
- [12] 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.
- [13] Donohue I, Hillebrand H, Montoya JM, et al. Navigating the complexity of ecological stability. *Ecol Lett* 2016;19:1172–85.
- [14] 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.
- [15] Durugbo II, 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.
- [16] Eren AM, Zozaya M, Taylor CM, Dowd SE, Martin DH, Ferris MJ. Exploring the diversity of *Gardnerella vaginalis* in the genitourinary tract microbiota of monogamous couples through subtle nucleotide variation. *PLoS One* 2011;6:e26732.
- [17] Fredricks DN, Fiedler TL, Marrazzo JM. Molecular identification of bacteria associated with bacterial vaginosis. *N Engl J Med* 2005;353:1899–911.
- [18] Fredricks DN. Molecular methods to describe the spectrum and dynamics of the vaginal microbiota. *Anaerobe* 2011;17(4):191–5.
- [19] Gajer P, Brotman RM, Bail G, et al. Temporal dynamics of the human vaginal microbiota. *Sci Transl Med* 2012;4(132):132ra52.
- [20] Garcia EM, Kraskauskienė V, Koblinski JE, 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).
- [21] Gibbs RS. Chorioamnionitis and bacterial vaginosis. *Am J Obstet Gynecol* 1993;169:460–2.
- [22] Gibbs RS. Asymptomatic bacterial vaginosis: is it time to treat?. *Am J Obstet Gynecol* 2007;196:495–6.
- [23] Gilbert NM, Lewis WG, 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.
- [24] Grady D, Thiemann C, Brockmann D. Robust classification of salient links in complex networks. *Nat Commun* 2012;3:864.
- [25] Guedou FA, Van Damme L, Mirembé 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.
- [26] Hickey RJ, Zhou X, Pierson JD, et al. Understanding vaginal microbiome complexity from an ecological perspective. *Transl Res* 2012;160:267–82.
- [27] Jakobsson T, Forsum U. *Lactobacillus iners*: a marker of changes in the vaginal flora?. *J Clin Microbiol* 2007;45:3145.
- [28] 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.
- [29] Klebanoff MA, Brotman RM. Treatment of bacterial vaginosis to prevent preterm birth. *Lancet* 2018;392(10160):2141–2.
- [30] 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.

- [31] Lambert JA, Kalra A, Dodge CT, et al. Novel PCR-Based Methods Enhance Characterization of Vaginal Microbiota in a Bacterial Vaginosis Patient before and after Treatment. *Appl Environ Microbiol* 2013;79(13):4181–5.
- [32] Lamont RF, Taylor-Robinson D. The role of bacterial vaginosis, aerobic vaginitis, abnormal vaginal flora and the risk of preterm birth. *BJOG* 2010;117:119–20.
- [33] Leitich H, Bodner-Adler B, Brunbauer M, et al. Bacterial vaginosis as a risk factor for preterm delivery: a meta-analysis. *Am J Obstet Gynecol* 2003;189(1):139–47.
- [34] Li W, Ma ZS. Dominance network analysis of the healthy human vaginal microbiome not dominated by *Lactobacillus* species. *Comput Struct Biotechnol J* 2020;18:3447–56.
- [35] Ling Z, Kong J, Liu F, et al. Molecular analysis of the diversity of vaginal microbiota associated with bacterial vaginosis. *BMC Genom* 2010;11:488.
- [36] Ma B, Forney LJ, Ravel J. The vaginal microbiome: rethinking health and disease. *Annu Rev Microbiol* 2012;66:371–89.
- [37] Ma ZS, Abdo Z, Forney L. Caring about trees in the forest: incorporating frailty in risk analysis for personalized medicine. *Personalized Med* 2011;8:681–8.
- [38] Ma ZS, Guan Q, Ye C, et al. Network analysis suggests a potentially 'evil' alliance of opportunistic pathogens inhibited by a cooperative network in human milk bacterial communities. *Sci Rep* 2015;5:8275.
- [39] Ma ZS, Ye DD. Trios—promising *in silico* biomarkers for differentiating the effect of disease on the human microbiome network. *Sci Rep* 2017;7:13259.
- [40] Ma ZS, Li LW. Quantifying the human vaginal community state types (CSTs) with the species specificity index. *PeerJ* 2017;5:e3366.
- [41] Ma ZS, Ellison AM. A unified concept of dominance applicable at both community and species scales. *Ecosphere* 2018;9(10):e02477.
- [42] Ma ZS, Ellison AM. Dominance network analysis provides a new framework for studying the diversity-stability relationship. *Ecol Monogr* 2019;89(2):e01358.
- [43] Ma ZS, Li LW, Gotelli NJ. Diversity-disease relationships and shared species analyses for human microbiome-associated diseases. *ISME J* 2019;13(8):1911–9.
- [44] Ma ZS. Critical network structures and medical ecology mechanisms underlying human microbiome-associated diseases. *iScience* 2020;23(6):101195.
- [45] Ma ZS. Testing the Anna Karenina Principle in human microbiome-associated diseases. *iScience* 2020;23(4):101007.
- [46] Ma ZS. Heterogeneity-disease relationship in the human microbiome associated diseases. *FEMS Microbiol Ecol* 2020;96(7):faa0933.
- [47] Mania-Pramanik J, Kerkar SC, Salvi VS. Bacterial vaginosis: a cause of infertility?. *Int J Std Aids* 2009;20:778–81.
- [48] Menard JP, Fenollar F, Henry M, et al. Molecular quantification of *Gardnerella vaginalis* and *Atopobium vaginae* loads to predict bacterial vaginosis. *Clin Infect Dis* 2008;47:33–43.
- [49] Muzny CA, Taylor CM, Swords WE, et al. An updated conceptual model on the pathogenesis of bacterial vaginosis. *J Infect Dis* 2019;220(8):1399–405.
- [50] Muzny CA, Blanchard E, Taylor CM, et al. Identification of key bacteria involved in the induction of incident bacterial vaginosis: a prospective study. *J Infect Dis* 2018;218:966–78.
- [51] Muzny CA, Kardas P. A narrative review of current challenges in the diagnosis and management of bacterial vaginosis. *Sex Transm Dis* 2020;47:441–6.
- [52] Muzny CA, Łaniewski P, Schwebke JR, et al. Host-vaginal microbiota interactions in the pathogenesis of bacterial vaginosis. *Curr Opin Infect Dis* 2020;33(1):59–65.
- [53] Onderdonk AB, Delaney ML, Fichorova RN. The human microbiome during bacterial vaginosis. *Clin Microbiol Rev* 2016;29(2):223–38.
- [54] Peebles K, Velloza J, Balkus JE, et al. High global burden and costs of bacterial vaginosis: a systematic review and meta-analysis. *Sex Transm Dis*. 2019;46(5):304–11.
- [55] Petrova MI, Reid G, Vanechoutte M, Lebeer S. *Lactobacillus iners*: friend or foe?. *Trends Microbiol* 2017;25(3):182–91.
- [56] Pleckaityte M, Zilnyte M, Zvirbliene A. Insights into the CRISPR/Cas system of *Gardnerella vaginalis*. *BMC Microbiol* 2012;12:301.
- [57] Plummer EL, Vodstrcil LA, Murray GL, et al. *Gardnerella vaginalis* clade distribution is associated with behavioural practices and Nugent Score in women who have sex with women. *J Infect Dis* 2020;221(3):454–63.
- [58] Randis TM, Ratner AJ. *Gardnerella* and *Prevotella*: co-conspirators in the pathogenesis of bacterial vaginosis. *J Infect Dis* 2019;220(7):1085–8.
- [59] Ravel J, Gajer P, Abdo Z, et al. Vaginal microbiome of reproductive-age women. *Proc US Natl Acad Sci* 2011;108(1):4680–7.
- [60] Ravel J, Moreno I, Simon C. Bacterial vaginosis and its association with infertility, endometritis, and pelvic inflammatory disease. *Am J Obstet Gynecol* 2020;224(3):251–7.
- [61] Ravel J, Brotman RM, Gajer P, et al. Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. *Microbiome* 2013;1:29.
- [62] Romero R, Hassan SS, Gajer P, et al. The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. *Microbiome* 2014;2:4.
- [63] Rosca AS, Castro J, Sousa LGV, et al. *Gardnerella* and vaginal health: the truth is out there. *FEMS Microbiol Rev* 2020;44:73–105.
- [64] Sabo MC, Richardson BA, Lavreys L, et al. Does bacterial vaginosis modify the effect of hormonal contraception on HIV seroconversion. *AIDS* 2019;33(7):1225–30.
- [65] Shekhtman LM, Bagrow JP, Brockmann D. Robustness of skeletons and salient features in networks. *J Complex Networks* 2014. <https://doi.org/10.1093/comnet/cnt019>.
- [66] Sobel JD. Is there a protective role for vaginal flora?. *Curr Infect Dis Rep* 1999;1:379–83.
- [67] Sobel JD, Kaur N, Woznicki NA, et al. Prognostic Indicators of Recurrence of Bacterial Vaginosis. *J Clin Microbiol* 2019;57(5).
- [68] Song Y, He L, Zhou F, et al. Segmentation, splitting, and classification of overlapping bacteria in microscope images for automatic bacterial vaginosis diagnosis. *IEEE J Biomed Health Inf* 2016;2016:1.
- [69] Spear GT, St John E, Zariffard MR. Bacterial vaginosis and human immunodeficiency virus infection. *AIDS Res Ther* 2007;4:25.
- [70] Srinivasan S, Liu C, Mitchell CM, et al. Temporal variability of human vaginal bacteria and relationship with bacterial vaginosis. *PLoS One* 2010;5:e10197.
- [71] Svare JA, Schmidt H, Hansen BB, et al. Bacterial vaginosis in a cohort of Danish pregnant women: prevalence and relationship with preterm delivery, low birthweight and perinatal infections. *BJOG* 2006;113(12):1419–25.
- [72] Taylor BD, Darville T, Haggerty CL. Does bacterial vaginosis cause pelvic inflammatory disease?. *Sex Transm Dis* 2013;40(2):117–22.
- [73] Torrone EA, Morrison CS, Chen PL, et al. Prevalence of sexually transmitted infections and bacterial vaginosis among women in sub-Saharan Africa: An individual participant data meta-analysis of 18 HIV prevention studies. *PLoS Med* 2018;15(2):e1002511.
- [74] US Preventive Services Task Force, Owens DK, Davidson KW, et al. Screening for bacterial vaginosis in pregnant persons to prevent preterm delivery: US preventive services task force recommendation statement. *JAMA* 2020;323(13):1286–92.
- [75] Vanechoutte M, Guschin A, Van Simaey L, et al. Emended description of *Gardnerella vaginalis* and description of *Gardnerella leopoldii* sp. nov., *Gardnerella pottii* sp. nov. and *Gardnerella swidsinskii* sp. nov., with delineation of 13 genomic species within the genus *Gardnerella*. *Int J Syst Evol Microbiol* 2019;69(3):679–87.
- [76] White BA, Creedon DJ, Nelson KE, et al. The vaginal microbiome in health and disease. *Trends Endocrinol Metab* 2011;22(1):389–93.
- [77] Fredricks DN, Fiedler TL, Thomas KK, et al. Targeted PCR for detection of vaginal bacteria associated with bacterial vaginosis. *J Clin Microbiol* 2007;45:3270–6.
- [78] Hilbert DW, Smith WL, Paulish-Miller TE, et al. Utilization of molecular methods to identify prognostic markers for recurrent bacterial vaginosis. *Diagn Microbiol Infect Dis* 2016;86:231–42.
- [79] Lev-Sagie A, Goldman-Wohl D, Cohen Y, et al. Vaginal microbiome transplantation in women with intractable bacterial vaginosis. *Nat. Med* 2019;25(10):1500–4.
- [80] Machado D, Castro J, Palmeira-de-Oliveira A, et al. Bacterial vaginosis biofilms: Challenges to current therapies and emerging solutions. *Front Microbiol* 2015;6:1528.
- [81] Meltzer MC, Desmond RA, Schwebke JR, et al. Association of *Mobiluncus curtisii* with recurrence of bacterial vaginosis. *Sex Transm Dis* 2008;35:611–3.
- [82] Verhelst R, Verstraelen H, Claeys G, et al. Cloning of 16S rRNA genes amplified from normal and disturbed vaginal microflora suggests a strong association between *Atopobium vaginae*, *Gardnerella vaginalis* and bacterial vaginosis. *BMC Microbiol* 2004;4:16.