Journal of
Medical Microbiology & Diagnosis

Liu et al., J Medical Microbiol Diagnosis 2012, S:5
DOI: 10.4172/2161-0703.S5-001

Research Article Open Access

Comparative Analysis of Vaginal Bacterial Diversity in Northern-Chinese Women Associated With or Without Bacterial Vaginosis

Yajuan Liu[1#], Wenbo Hao[1#], Chaohui Duan[2#], Jianping Tan[3], Weiwen Xu[1], Shimeng Zhang[1], Xiaoqing Liao[1], Ming Li[1*], Shuhong Luo[1*]

[1]Institute of Antibody Engineering, School of Biotechnology, Southern Medical University, 1838 N. Guangzhou Avenue, Guangzhou People's Republic of China
[2]Laboratory of Clinical Immunology, the Sun Yat-Sen Memorial hospital, Sun Yat-Sen University, Guangzhou People's Republic of China
[3]Department of obstetrics and gynecology, the Sun Yat-Sen Memorial hospital, Sun Yat-Sen University, Guangzhou People's Republic of China
[#]These authors contributed equally to this work.

Abstract

Background: Bacterial Vaginosis (BV) is the common vaginal infection in women and it has been linked to enhanced risks for pre-term birth, pelvic inflammatory disease and sexually transmitted diseases. BV is caused by a disorder of vaginal microbiota which changes from the normal lactobacillus dominated community to a more diverse community of non-lactobacillus bacteria. Several previous reports analyzed the overall vaginal microbial communities of volunteers from limited sampling area and they suggested a possible link between the vaginal microbial contents and the ethnicity of women. Here, we analyzed the diversity of vaginal microbiota in 10 subjects associated with BV (BV+) and 10 subjects without BV (BV-) from the metropolitan area of Herbing in Northern China using full-length 16S rDNA.

Results: The vaginal bacterial communities detected in subjects with BV were much more taxon rich and diverse than those without BV. At a 97% sequence similarity cutoff, the number of operational taxonomic units (OTUs) per 10 subjects with BV was nearly three times greater than 10 subjects without BV by 29.4 ± 9.3 versus 11.7 ± 7.8 (Mean ± SD). Our data confirmed that there is a shift in the abundance of bacterial species present in the vaginal environment when BV and non-BV groups were compared. Each sequence read was assigned to a genus or a species when possible. Principal component analysis was performed at genus levels. Most BV+ samples clustered together while there were two clusters among BV- samples. Several bacteria have been found to be associated with BV, including *Gardnerella, Atopobium, Peptoniphilus, Leptotrichia/Sneathia, Prevotella, Parvimonas* and *Dialister,* Based on result of classification, four possible novel phylotype microorganisms were found.

Conclusions: The data presented here on the composition and richness of the vaginal microbial ecosystem in BV and health state will provide the depth insight in the etiology of BV.

Keywords: Bacterial vaginosis; 16S rDNA amplification; Clone library; Vaginal flora; Microbiology

Abbreviations: BV: Bacterial vaginosis; OTUs: Operational taxonomic units; PCA: Principal component analysis

Background

Bacterial vaginosis (BV) is the common vaginal infection in women of reproductive age, and is the most common etiology of vaginal symptoms which prompt women to seek medical care [1]. Numerous health problems including preterm labor resulting in low birth weight [1], pelvic inflammatory disease [2], and acquisition of HIV [3] are closely related to BV.

BV is caused by an imbalance of naturally occurring bacterial flora in healthy vagina with the common normal bacterial flora of *Lactobacillus crispatus* and *Lactobacillus iners. Lactobacilli* include some hydrogen peroxide-producing species that help to prevent other vaginal microorganisms from excessive growth. Previous reports showed that BV patients have dromatic changes in vaginal from the normal lactobacillus dominating community to the diversity of non-lactobacillus bacteria [4]. Alone the absence of *Lactobacillus* does not define an unhealthy state. Complementarily, the presence of solely, or a combination of, *Atopobium, Gardnerella, Peptostreptococcus, Prevotella, Pseudomonas*, and/or Streptococcus (often noxious bacteria when in/on humans) does not define an unhealthy state [5]. In the past, the composition and diversity of human vagina microbiota were detected by cultural methods which had many limitations for the really circumstance. The advent of PCR based techniques and pyrosequencing made it possible to further examine this complex microbial niche. The small ribosomal subunit or 16S rDNA gene was often used as the most common target for molecular identification of bacteria. Bacterial 16S rDNA gene is present in all bacteria and has conservative regions that can be targeted with broad range PCR primers [6]. Areas of amplified sequences can be used to characterize and identify the bacteria origin, in order to determine their phylogenetic relationships.

In this study, we used the full-length 16S rDNA to determine the diversity of vaginal microbiota in 10 BV+ and 10 BV- women from Haerbin in Northern China and to get the basic data for the bacterial imbalance when BV occured. Our results revealed that analysis of vaginal microbiota and the clustering of microbial profile could help in the identification of bacterial vaginosis and assessment of women's vaginal health.

***Corresponding author:** Dr. Shuhong Luo, Institute of Antibody Engineering, School of Biotechnology, Southern Medical University, Guangzhou Avenue, Guangzhou, Guangdong Province, China; Tel: 8620-61648325 ; Fax : 8620-61648554; E-mail: shluo815@yahoo.com

Received March 17, 2012; **Accepted** May 22, 2012; **Published** May 30, 2012

Citation: Liu Y, Hao W, Duan C, Tan J, Xu W, et al. (2012) Comparative Analysis of Vaginal Bacterial Diversity in Northern-Chinese Women Associated With or Without Bacterial Vaginosis. J Medical Microbiol Diagnosis S5:001. doi:10.4172/2161-0703.S5-001

Methods

Human subjects

This study was approved by both the ethics committee of Southern Medical University and by the Heilongjiang Provincial People's Hospital. Twenty women were recruited for this study. All the women enrolled provided informed consent and none of them declined participation. The study was supervised by experts from Southern Medical University and Heilongjiang Provincial People's Hospital. These subjects were at various (recorded) times of their menstrual cycle with 19-55 years of age, including 10 BV positive (BV+) women (BV+ group, aged 36.8 ± 15.02.) and 10 BV negative (BV-) women (BV- group, aged 29.1 ± 4.43), who came to the Department of Obstetrics and Gynecology, Heilongjiang Provincial People's Hospital, for routine gynecology examination from August 2009 to July 2011.

BV status was assessed based on Amsel criteria for all subjects [7]. The participants who met three or more of the following criteria were clinically diagnosed as BV+: homogeneous vaginal discharge, >20% clue cells on wet mount microscopy, elevated pH (≥ 4.5) of vaginal fluid, and release of a fishy amine odor on addition of 10% potassium hydroxide solution to vaginal fluid ("whiff" test). Diagnosis was confirmed using BV Quick-Test sialidase test kit (Zhuhai Livzon Pharmaceutical Group Inc, China, Cat. No.2401325). Participants without these symptoms were defined as BV negative group (BV-).

Participants were excluded by any of the following exclusion criteria : <18 years of age, pregnancy, taking contraceptive steroids, complaints of urogenital symptoms or noticeable infection on physical examination of the urogenital tract, diabetes mellitus, the use of antibiotics or vaginal antimicrobials in the previous month, presence of an intrauterine device, vaginal intercourse within the latest 3 days, known active infection due to Chlamydia, yeast, *Trichomonas vaginalis,* clinically apparent herpes simplex infection, or defined diagnosed HPV, HSV-2, or HIV-1 infection. The clinical data of each participant were summarized in Table 1.

Sample No	Age	BV Test[a]	Clue cell	rod-shaped bacteria	WBC Count[b]	Clinical Diagnosis[c]
BV+ Group						
124	30	+	+	++	W++	BV
127	27	+	+	+	W+++	BV
130	55	+	-	++	W++++	BV
131	19	+	+	++	W++	BV
133	30	+	+	++	W++++	BV
227	57	+	+	++	W+++	BV
228	58	+	+	++	W++	BV
232	29	+	+	+	W+++	BV
245	21	+	+	++	W++++	BV
251	42	+	+	+	W++	BV
BV- Group						
113	29	-	-	Miscellaneous+	W+	PE
125	21	-	-	Miscellaneous+ +	W++	PE
136	28	-	-	Miscellaneous+	W++	PE
138	30	-	-	Miscellaneous+	W++	PE
146	32	-	-	Miscellaneous+	W++	PE
223	22	-	-	Miscellaneous+	W++	PE
229	32	-	-	Miscellaneous+	W++	PE
235	31	-	-	Miscellaneous+	W++	PE
237	35	-	-	Miscellaneous+	W+	PE
249	31	-	-		W+	PE

[a]BV test is based on a sialidase test kit.
[b] W values are the white blood cell counts per viewing field under microscope. W+ indicates a value greater than 10 and W++, W+++, W++++ greater than 20, 30, and 40 respectively.
[c]Diagnosis was based on physician's comments. PE indicates that the sample was collected during a routine physical examination.

Table 1: Patient information of volunteer subjects.

Sample collection and preparation

The sample swabs were collected as described by Kin et al. [8]. Two swabs were taken from each woman. One was used for sialidase test, vaginal smears, and saline microscopy, the other was used for bacterial genomic DNA extraction. This vaginal swab was placed in sterile cryovial with 1 ml normal saline, packaged in ice packs and transferred to the laboratory immediately in ice-box, and stored at -80 °C for future test.

Extraction of genomic DNA

The bacterial cells retrieved on swabs were vigorously agitated to dislodge cells. The cells were pelleted by centrifugation at 10,000 g for 10 min. Then bacterial genomic DNA was extracted using QIAamp DNeasy Blood and Tissue Kit (QIAGEN, Hilden, Germany; Cat. No. 69504) followed the manufacturer′s instructions. Briefly, the bacterial pellet was suspended in 200 µl of normal saline, and then added 20 µl of proteinase K solution and 200 µl of lysis buffer (Buffer AL). Vertex the mixtures for homogenization and incubated at 56 °C for 30 min. Then, 200 µl of 100% ethanol was added and mixed thoroughly, the mixture was then pipeted into the QIAamp spin column and centrifuged at 8,000 g for 1 min. The QIAamp spin column was placed in a new 2 ml collection micro-tube, and the containing filtrate was discarded. The spin columns were washed with 500 µl buffer AW1 and AW2 provided in the kit. Finally, the genomic DNA was eluted with 100 µl of elution buffer (Buffer AE). The concentration of DNA was determined by NanoDrop ND-1000 spectrophotometer (Thermo Electron Corporation, USA). The integrity and size of the extracted DNA were estimated and checked by 1.0% agarose gel electrophoresis containing 0.5 mg/ml ethidium bromide. All the DNA samples were stored at -20 °C until use.

PCR amplification

Highly conserved sequences of the *Escherichia coli* 16S rDNA was used to design the primers for PCR amplication. The forward primer was 8f: 5′-AGAGTTTGAT CMTGGCTCAG-3′, M = A + C. In the European Ribosomal RNA Database, 8f is called BSF8/20 (www.psb.ugent.be/rRNA). The reverse primer was 1492r: 5′-TACG GYTACCTTGTTACGACTT-3′, Y = C + T [9,10]. The anticipated amplification product from the of microbial genomic DNA templates was nearly complete 16S rDNA (>1.4 kb) [11]. PCR amplification was performed with Promega (GoTaq, Madison WI) PCR kit and followed the manufacturer′s instructions. Total 50 µl PCR mixture contained 0.1 µM of each primer, 25 µl of Promega master mix, 3 µl of DNA extract and distilled water. The amplify procedure included 95°C for 3 min, amplified for 30 cycles where each cycle consisted of 95°C, 1 min; 52 °C, 1min; and 72°C, 2 min 30 sec, followed by 72°C, 10 min. The 1.4 kb amplicons were purified used the gel extraction kit (Omega Cat. D2500).

16S rDNA gene cloning library construction and sequencing

16S rDNA PCR products were cloned into the pMD19T (Takara, Dalian) vector and transferred to *E. coli* TOP 10 competent cells (TianGen) according to the manufacturer's recommendations. The TOP 10 cells were plated onto Luria–Bertani agar plates containing

100 μg/ml ampicillin and 20 μg/ml X-gal and incubated overnight at 37°C. Approximately 200 white, well-isolated colonies were randomly selected from each of library. Single colonies were picked from the agar plates and transferred with sterile tips to the wells of a 96 well plate filled with LB Broth supplemented with 200 μg ampicillin/ml and incubated at 37°C for 3 h. For each sample, at least one 96-well plate containing a library with >90% of the colonies carrying cloned inserts was sequenced by using Applied Biosystems BigDye-terminator chemistry and the Applied Biosystems 3730 DNA Analyzer (Carlsbad, CA) with universal primer M13 reverse -48.

Sequence processing and data analysis

The clone sequences were edited and processed by soft package. Raw sequence data were base-calling by PHRED and assigns a quality score as Q20 to each base edited and processed to accurate reads for alignment [12,13]. In all cases, vector and primer sequences were removed and quality trimmed from the clone sequences by using the LUCY program [14,15]. 37 clones of sequences are shorter than 200 bps or empty vector without insert sequence, thus were cut off. The related bacterial 16S rDNA sequences in the RDP II database (http://rdp.cme.msu.edu/) with high Seqmatch scores (>0.9), were selected and used to assign by RDP naïve Bayesian rRNA Classifier for further analysis [15,16]. Analysis of the 8 clone sequences, which did not match, revealed that nearly all of them were probably not unique or unknown phylotypes. Based on the Seqmatch results, the clone sequences could be assigned to the type strains in the RDP II.

Statistical analysis

Bray-Curtis distance analysis was used to determine the differences in microbial population profiles. The Bray-Curtis dissimilarity index [17,18], which is equivalent to a doubly weighted form of the Sørensen-Dice dissimilarity index, was calculated according to the equation when Σ_{ixij} =1 and Σ_{ixik} =1, where d_{jk} is the dissimilarity or distance between communities j and k, and x_{ij} and x_{ik} are the fractional or percentage populations of component i found in communities j and k. This distance was used to calculate the distance matrix for cluster analysis using the unweighted-pair group method using average linkages in the PHYLIP package. Phylotype diversity of the combined vaginal libraries within each subject was estimated by Simpson's reciprocal and Shannon-Wiener's diversity indices, using the EstimateS software (http://viceroy.eeb.uconn.edu/estimates). The OTU, rarefaction, Chao1 and ACE estimation were analyzed using the mothur [19] (v.1.6.0, http://www.mothur.org/wiki/Main-Page). The statistical significance of differences in microbial community composition, and Shannon and Simpson index between sample categories was determined by SPSS with chi-test and t-test. For bootstrap values, three methods (Fitch, parsimony, and maximum likelihood) with 100 re-samplings were performed using FITCH, DNAPARS, DNAMLK, SEQBOOT, and CONSENSE in PHYLIP.

Results

Sequence analysis

Clones were sequenced with a single read using M13 rev-48 primer and showed that nearly complete 16S rDNA PCR product was cloned into pMD19T vector. Total of 3245 sequences passed quality control and were used for data analysis; 1893 sequences for BV+ group and 1346 sequences for BV- group. Sequences were classified based on self-similarities rather than matches to an external database and were grouped into operational taxonomic units (OTUs) with cutoffs of each sequence similarity using the MOTHUR package (http://www.mothur.org/wiki/Main-Page) [19]. As described previously [20], rarefaction curves were generated for unique, 1%, 2%, 3%, sequence dissimilarities. For phylogenetic analyses, the core data set of representative sequences with 97% 16S rDNA sequence similarity to definite the phylogenetic species. As shown in Figure 1, the rarefaction curve in BV- group almost reached the saturation level. However, there was much more richness in bacterial diversity in BV+ individuals than in BV- individuals at the 97% similarity level. The number of OTUs continued to increase at the 97% or 98% similarity level in BV+ group (Figure 1), which indicated that additional sampling was needed to determine the true microbial diversity in BV+ vaginal community.

The taxonomic assignments of vaginal bacterial community members and the associated diversity estimation for each subject are shown in Table 2. The sequences identified in these groups covered the majority of the clone libraries presented in our detected samples for the Good's coverage weremore than 98.0% for all sequences in the two groups according to the estimated species diversity indices. Furthermore, as estimated by Chao1 and ACE diversity indices, the number of OTUs detected was close to the total number of OTUs for the two groups of vaginal communities analyzed at the 3% dissimilarity level. Based on the number of sequences or OTUs and relative proportions, Chao1 and ACE richness estimators were measured for the number of microbial species in the samples. This also proved that our clone libraries were well-covered during sequencing. Taxon definitions of 99%, 98%, and 97% OTUs were assigned using the Mothur package. The numbers of OTU taxa per subjects with BV of Mean ± SD were 39.8 ± 11.9, 31.2 ± 11.1, 29.4 ± 9.3 versus 16.2 ± 11.6, 11.3 ± 8.5, 11.7 ± 7.8 correspond to the subjects without BV at 99%, 98%, 97% similarity respectively. The OUT taxa detected per subject were dramatically higher ($P \leq 0.001$) in subjects with BV than in subjects without BV for all phylotype definitions (Figure 2). Similarly,

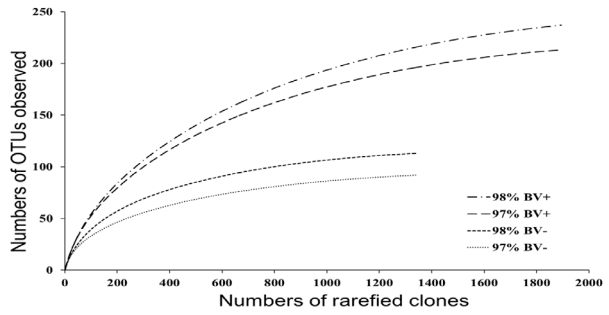

Figure 1: Rarefaction curves for BV+ and BV- groups used to estimate richness at a 97% or 98% similarity level.

Group	Reads	OTUs[1]	Good[2]	ACE	Chao	Shannon[3]	Evenness	Simpson
BV+	1893	213	0.981	230	222	4.469	0.592	0.02544
BV-	1346	92	0.989	99	96	3.401	0.472	0.068199

[1]The operational taxonomic units (OTUs) were defined with 3% dissimilarity level.
[2]The coverage percentage (Good), richness estimators (ACE and Chao1) and diversity indices (Shannon and Simpson) were calculated using Good's method. (Good, 1953) and the MOTHUR program.
[3]The Shannon index of evenness was calculated with the formula E = H/ln(S), where H is the Shannon diversity index and S is the total number of sequences in that group.

Table 2: Comparison of phylotype coverage and diversity estimation of the 16S rRNA gene libraries at the 3% dissimilarity.

diversity was higher (P <0.005) in subjects with BV than that without BV. The Shannon diversity index in subjects with BV was 2.79 ± 0.5, and in subjects without BV was 1.56 ± 0.9 based on the 97% similarity taxonomic definition.

The Shannon and Simpson indices [21], for both clone libraries were far from the critical values, Shannon Index (3.401 in BV- group *Vs.* 4.469 in BV+ group) and Simpson Index (0.068199 in BV- group *Vs.*0.02544 in BV+ group). The diversity of both groups was abundant; however the diversity was higher in subjects with BV, which the Shannon diversity index was 1.3 times greater than subjects without BV. The vaginal community in BV+ group (Evenness = 0.592) was more even than in BV- group (Evenness=0.472) (Table 2) indicating that bacterial community in BV+ group had greater species diversity.

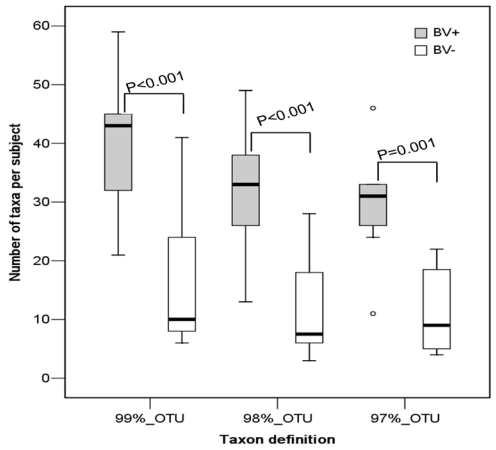

Figure 2: Numbers of taxa per subject using three different taxon definitions. Taxon definitions of 99%, 98%, and 97% OTUs were assigned using the Mothur package.
The shaded boxes encompass the 25th to 75th percentiles of the data, the boldface lines indicate means, and the whiskers span the 5th to 95th percentiles. The asterisks indicate significantly (P<0.001) differences in mean values for subjects with BV versus subjects without BV, as determined by t tests.

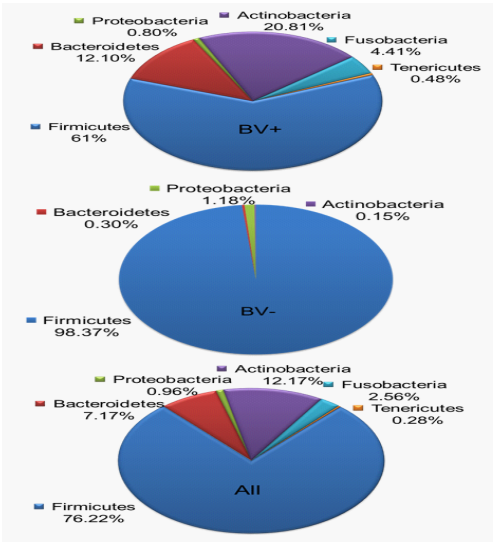

Figure 3: Pie chart representation for the overall abundance of phyla among the 20 libraries.

Differential bacterial abundance found in BV+ and BV- Groups

A significant difference (P <0.001) was found between the subjects with BV and those without BV according to the mean number of taxa in each sample of Mean ± SD with the 11.8 ± 4.44 in BV+ group and 3.9 ± 3.35 in BV- group.

The overall abundance was presented as pie charts in Figure 3. At phylum level, *Firmicutes* was the most abundant member in BV- group and *Firmicutes* and *Actinobacteria* was more abundant in BV+ group (Figure 3A and B). *Fusobacteria* and *Tenericutes* were only found in samples from subjects with BV. There were differences in the relative proportions of major groups of bacteria for the taxonomic distribution of phylotypes. Taxonomic affiliations and relative proportions of sequences were summarized at the genus (Figure 4A) and phylum (Figure 4B) levels.

Proportions of *Actinobacteria*, and *Bacteroidetes* were 140 times and 40 times higher respectively in subjects with BV than those without BV. However *Firmicutes* in subjects with BV was only 0.62 times of that without BV. The dominant microflora is *Lactobacillus iners* in subjects without BV, which accounted for 83.73% of all bacterial communities (Figure 4A). In contrast, subjects with BV did not possess a single dominant taxon but rather harbored a diverse array of vaginal bacteria, many present at low relative abundance, including *Gardnerella, Atopobium, Prevotella, Gemella, Aerococcus, L.iners, Anaerococcus, Parvimonas, Peptoniphilus, Megasphaera, Dialiste, and Sneathia*. From the above, we can see that BV+ group has a higher degree of microbial diversity.

The abundance of bacterial phylotypes from each subject was summarized in Table 3. Several species showed preferential occurrence in subjects of BV+ group with t-test p value < 0.1 and this preference was analyzed by chi-test according to presence of species in the subject (Table 4). Most of these species were more prevalent in BV+ group.

Subjects	No. of Sequences	Phylum					
		Firmicutes	Bacteroidetes	Proteobacteria	Actinobacteria	Fusobacteria	Tenericutes
BV+ Group							
124	165	38.55	3.01	0	51.81	6.02	0.6
127	153	80.39	5.88	3.92	5.23	4.58	0
130	121	5.79	5.79	0	28.1	21.49	0
131	190	48.42	1.05	0	50	0	0.53
133	181	66.3	10.5	0	15.47	7.73	0
227	231	75.32	16.88	0	2.6	5.19	0
228	229	94.76	1.75	0	3.49	0	0
232	200	48	22	0.5	25	4.5	0
245	224	78.13	0	3.57	15.18	0	3.13
251	190	47.37	37.89	0	12.11	2.63	0
BV- Group							
113	118	100	0	0	0	0	0
125	244	99.59	0.41	0	0	0	0
136	199	99.5	0	0.5	0	0	0
138	117	99.15	0	0.85	0	0	0
146	153	89.54	0.65	8.5	1.31	0	0
223	101	100	0	0	0	0	0
229	106	100	0	0	0	0	0
235	105	97.14	1.9	0.95	0	0	0
237	105	100	0	0	0	0	0
249	104	100	0	0	0	0	0

Table 3: Abundance of bacterial phylotypes at Phylum level.

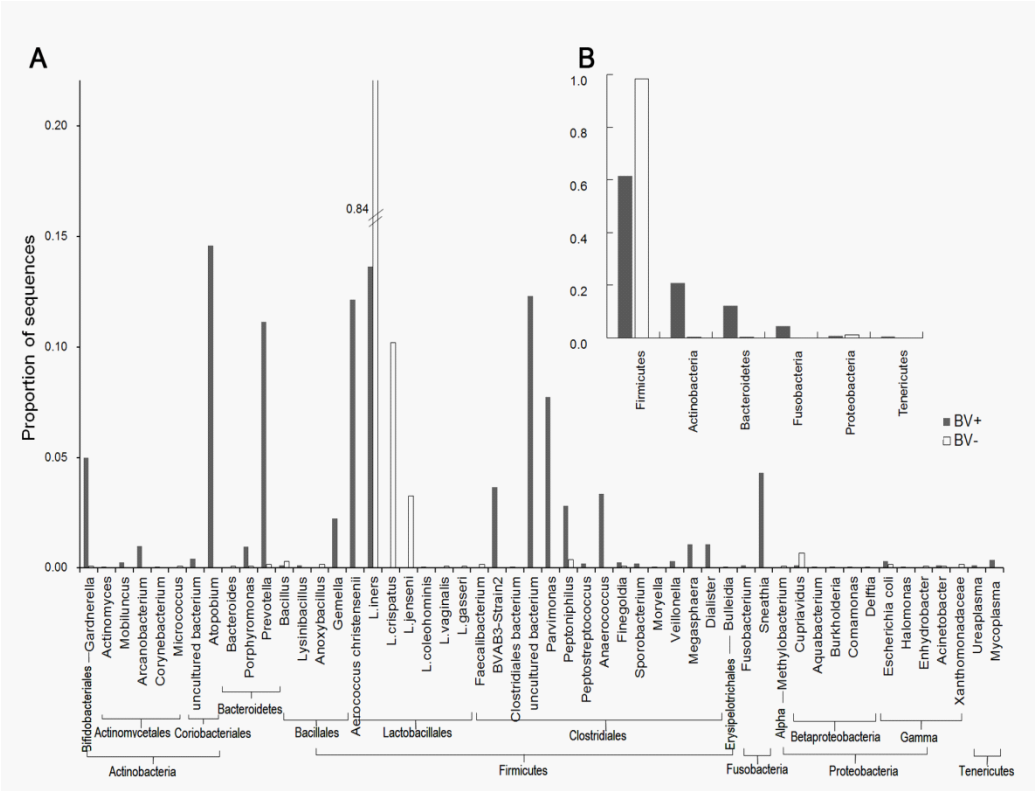

Figure 4: Taxonomic affiliations and relative proportions of sequences summarized at the genus (A) and phylum (B) levels. The taxa in panel A represent the closest matches from the RDP database. The y axes in both panels show the proportion of sequences from subjects with and without BV calculated separately. Note the break in the axis for *Lactobacillus iners* sequences from subjects without BV, where *lactobacilli* comprise 84% of the sequences.

Gardnerella vaginalis was one of the bacteria first found to be associated with BV [22]. In our research, *Gardnerella vaginalis* was found in 4 subjects in BV+ group, with the occurrence of 10.3%, 17.37%, 0.87% and 15.18% in the subject 124, 131,228 and 245 respectively. In BV- group, *Gardnerella vaginalis* was found in only one subject 146 with the occurrence of 0.65%. In this study, the t-test p value for abundance was only < 0.1 and chi-test p value for presence was < 0.5. Although *Gardnerella vaginalis* was not more prevalent in BV+ group, its presence was more common in BV+ group (Figure 4).

There was a significant difference between BV+ and BV- group in almost genera obtained from the vagina (Table 4). Not surprisingly, the *Atopobium, Prevotella, Sneathia* (P<0.01 for chi-test, and P<0.05 for t-test) had significantly strong association with BV. Among these predominant genera, *Gemella, Anaerococcus, Parvimonas, Peptoniphilus, Dialiste, Megasphaera* had weak associations with BV (P<0.05 for only one test). Although they had been detected by previously studies, they were not identified as BV-associated bacteria [5,23-26].

Lactobacillus bacteria were more abundant in BV- group, consisting of 97.6% ± 9.6% and significantly higher than that in BV+ group, which was 12.2% ± 26.6% (P<0.001) (Figure 4). Among the *Lactobacillus* species, *L. iners, L. jensenii, L. crispatus, L. vaginalis* and *L. gasseri* were found with overall occurrence of 83.73%, 3.25%, 10.21%, 0.07% and 0.07% in BV- group. Only *L. iners* and *L. coleohominis* were found with 13.63% and 0.05% in BV + group. *L. iners* was more prevalent in BV- group with t-test p value of 0.001 and chi-test p value of 0.001(Table 3). Chinese women are more likely to be colonized with *L. iners* than other Lactobacillus species in the vaginal ecosystem. This is consistent with Ravel, that among 96 Asian women, 41 had a vaginal microbiota dominated by *L. iners* [26]. The diversity of vaginal microbiota in healthy women from the Haerbing, Northern China, compared with the other area of China investigated [27,28]. *L. vaginalis* was present in subject 125 with 0.41% prevalence, and *L. gasseri* in subject 229 with 0.94%.

Based on results of classification, four possible novel phylotype microorganisms were found. Two uncultured bacterium classified into the Ruminococcaceae family were apparent to be associate with BV, that uncultured bacterium AY958888 showed significant BV-associate (P<0.05 for chi-test and P<0.1 for t-test), another uncultured Acetivibrio bacterium (P<0.1 for chi-test) which was indifined as BVAB3-Strain 2 by Fredricks [23].

Two additional unidentified bacteria were also found in this study, although not associated with BV. One classified into Coriobacteriaceae bacterium was presented in subject 130, 133 and 232 with prevalence of 1.65%, 0.55% and 2.5%. The other classified into Clostridiales bacterium was presented in subject 131 with 0.52%.

Principal component analysis

Principal Component Analysis (PCA) on the genera composition of microbes in each subject that represents an environment was performed using Canoco (Version 4.51) [29]. The clustering of the samples and the vectors contributing to the spread of the data points

			BV+		BV-		P value	
			Mean (%)	std	Mean (%)	std		
Genera	Genbank No.	prevalence (%)					T-test	chi-test
Gardnerella	M58744.1	2.688	4.366	7.048	0.065	0.207	0.086	0.218
Actinomyces	HQ850579.1	0.031	0.065	0.207	0	0	0.343	0.739
Mobiluncus	AJ427624.2	0.154	0.249	0.486	0	0	0.14	0.28
Arcanobacteriu	HQ712123.1	0.494	0.884	2.795	0	0	0.343	0.739
Corynebacterium	AJ277970.1	0.031	0.053	0.166	0	0	0.343	0.739
Micrococcus		0.031	0	0	0.065	0.207	0.343	0.739
Uncultured	AY959023.1	0.247	0.471	0.887	0	0	0.128	0.28
Atopobium	AF325325.1	8.496	16.464	16.817	0	0	0.013	0
Bacteroides	AB510708.1	0.031	0	0	0.041	0.13	0.343	0.739
Porphyromonas	AB547668.1	0.587	1.017	1.771	0.065	0.207	0.125	0.143
Prevotella	AB547706.1	6.549	11.69	12.438	0.19	0.602	0.017	0
Lysinibacillus	JF309274.1	0.062	0.105	0.333	0	0	0.343	0.739
Anoxybacillus	AJ551330	0.062	0	0	0.131	0.413	0.343	0.739
Gemella	EU427463.1	1.297	2.493	4.311	0	0	0.101	0.002
Aerococcus	Y17318.1	7.074	10.249	29.372	0	0	0.298	0.481
L. iners	AY526083.1	42.91	12.159	26.198	81.476	29.976	0	0
L.crispatus	Y17362	4.263	0	0	12.276	30.624	0.237	0.023
Sporobacterium	EU483154.1	0.124	0.211	0.666	0	0	0.343	0.739
Moryella ;	AF527773.1	0.031	0.053	0.166	0	0	0.343	0.739
Anaerococcus	GQ422749.1	1.946	3.879	7.537	0	0	0.138	0.023
Finegoldia	AP008971.1	0.185	0.278	0.593	0.065	0.207	0.306	0.684
Parvimonas	GU470891.1	4.51	8.182	10.709	0	0	0.039	0.063
Peptoniphilus	AY244779.1	1.792	3.29	5	0.327	1.033	0.097	0.009
Peptostreptococcus	AY326462.1	0.124	0.223	0.382	0	0	0.098	0.28
Faecalibacterium	AF371731	0.062	0	0	0.082	0.259	0.343	0.739
Acetivibrio BVAB3	GQ900632	2.132	3.508	6.141	0	0	0.104	0.28
Clostridiales bacterium	AF481208.1	0.031	0.053	0.166	0	0	0.343	0.739
uncultured bacterium	AY958888	7.167	10.973	18.509	0	0	0.094	0.063
Veillonella	AY244769.1	0.185	0.361	1.143	0	0	0.343	0.739
Megasphaera	DQ666098	0.618	1.044	1.811	0	0	0.102	0.143
Dialister	AB626633.1	0.618	1.056	1.45	0	0	0.047	0.739
Bulleidia		0.031	0.043	0.137	0	0	0.343	0.739
Fusobacterium	GQ301043.1	0.062	0.109	0.235	0	0	0.178	0.481
Sneathia	GQ179730.1	2.502	5.106	6.345	0	0	0.031	0.007
Methylobacteriu	GU294335.1	0.031	0	0	0.085	0.27	0.343	0.739
Cupriavidus	AB109753.1	0.34	0.131	0.413	0.588	1.86	0.458	0.971
Aquabacterium	AJ744886	0.031	0.05	0.158	0	0	0.343	0.739
Burkholderia		0.031	0.045	0.141	0	0	0.343	0.739
Comamonas	AB277849.1	0.031	0.065	0.207	0	0	0.343	0.739
Mycoplasma		0.216	0.313	0.988	0	0	0.343	0.739
Delftia	AB074256.1	0.031	0.065	0.207	0	0	0.343	0.739
Escherichia coli;	AF233451	0.247	0.268	0.847	0.131	0.413	0.651	0.971
Halomonas		0.031	0.045	0.141	0	0	0.343	0.739
Enhydrobacter	AJ550856	0.031	0	0	0.065	0.207	0.343	0.739
Acinetobacter	AJ275041.2	0.093	0.131	0.413	0.095	0.301	0.829	0.971
Xanthomonadaceae	GQ926873.1	0.062	0	0	0.116	0.246	0.172	0.481
Ureaplasma	AF073459.1	0.062	0.113	0.239	0	0	0.169	0.481

[a]It is usually assumed that the prevalence distributions of parameters are Gaussian distribution. The parameters statistical test was performed by independent Student's t test first and

[b]For those Equal variances not assumed, chi-square test was used to calculate the P value

The underscore characters for the P value<0.05, indicated that the prevalence of BV positive group is significantly differed from BV negative group.

Table 4: Genera of bacteria with high overall prevalence or with differential occurrence in BV+ vs BV- group.

were showed by the biplot in Figure 5. *Lactobacillus, Anaerococcus, Prevotella* and *Sneathia* contributed to the formation of a cluster, respectively. *Parvimonas, Atopobium, Gemella, Gardnerella,* and *Dialister* altogether contributed to the separation of a cluster. When the clusters from the PCA results were compared with the sialidase based BV test, nine out of ten subjects in BV+ group fall into the same cluster. The only exception was subject 245 with the dominant consist of *L. iners*. Although subject 245 had been treated for BV, the sialidase

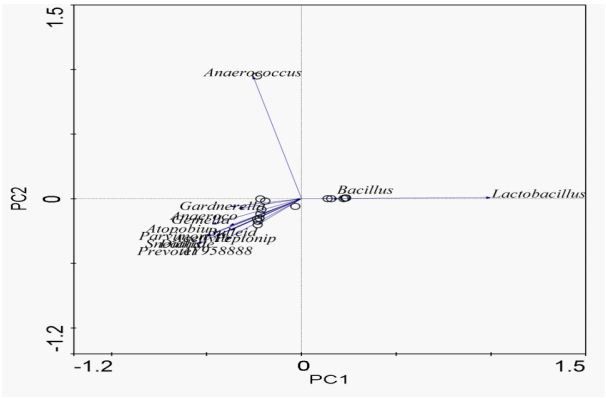

Figure 5: Biplot of principal component analysis results for the micobial centents at genus binning level.

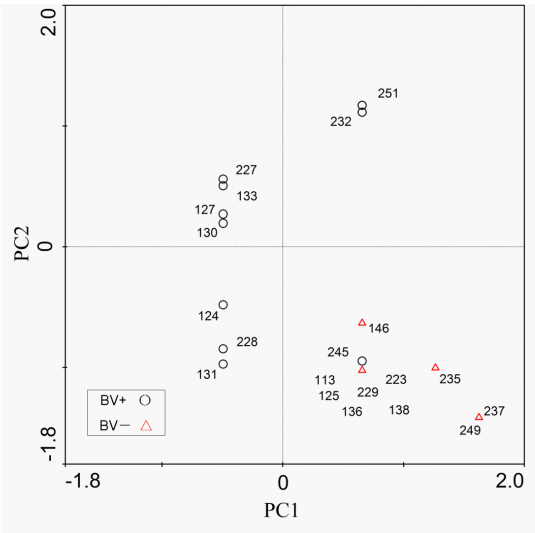

Figure 6: Clustering of subjects after principal component analysis.

test still showed positive result (Figure 6). This suggests that absence of BV associated bacteria from the survey of the vaginal microbiota is not associated with the immunological response to a recent BV episode.

Discussion

The human vaginal microbial ecosystem plays an important protective role in maintaining women's health. Acquisition of the diversity and composition of the human vaginal microbial ecosystem in health and diseased state can be valuable in understanding the etiology of the disease and to prevent and control them. Here, we analyzed the diversity of vaginal microbiota in BV positive and healthy women using full-length 16S rDNA to better understand the diversity of healthy and diseased vaginal bacterial communities.

In subjects with clinically defined BV, the apparent change was the abundance of bacterial species and BV is associated with a dramatic increase in the taxonomic richness and diversity of the vaginal bacterial community. As to the bacterial cluster, there are 9 in BV subjects but only 1 in health subjects. Several bacteria such as *Atopobium, Prevotella, Parvimonas, Peptoniphilus, Dialiste,* and *Sneathia* were found to be highly associated with BV. We also found that the presence of BV-associated organisms can also be detected in women with normal flora..

In the previous reports, the dominating vaginal microbiota was *Lactobacilli* [27,28,30-32] .The species *Lactobacillus iners, Lactobacillus crispatus, Lactobacillus jensenii*, and *Lactobacillus gasseri* were found in the Asian Japanese, white and black women [30-33] while only the *Lactobacillus iners, Lactobacillus crispatus* were found in Eastern Chinese (Shanghai and Zhejiang) women [27,28]. In our study, the diversity of vaginal microbiota in healthy women from the Haerbing, Northern China was investigated. *L. vaginalis* was found in this investigation. Among the *Lactobacillus* species, *L. iners, L. crispatus, L. jensenii, L. vaginalis, L. gasseri and L.coleohominis* were found with overall occurrence of 42.91%, 4.26%, 1.35%, 0.03%, 0.03% and 0.03% respectively. *L. iners* and *L.crispatus* were more prevalent with t-test p value of 0.000, 0.237 and chi-test p value of 0.000, 0.023 (Table 4, Figure 4). Our results showed that the vaginal bacterial communities of healthy women are not always dominated by *Lactobacillus* species. The difference may be caused by the host genetic factors, cultural, behavioral and environmental differences.

In our research, we observed total of 213 taxa across all subjects with BV when taxon defined as 97% OTU definition, but richness estimates of the true number of taxa present ranged from 217 to 236 (95% confidence interval), with an average estimate of 222 taxa (Table 2). Maybe the more samples could really reveal even greater richness and diversity in subjects with BV. While to subjects without BV, our samples were enough because we got 92 taxa with a 97% OTU definition with the total ranged from 93 to 107 (95% confidence interval) (Table 2).

The variability in community membership and structure has important implications for understanding the etiology of BV and for developing diagnostic tools. Our data revealed that the vaginal bacterial communities found in subjects with and without BV are distinctly different. Notably, the structures of these vaginal bacterial communities have high inter subject variability within each clinical group. This is consistent with the reported inter subject variability in the vagina [23] and other areas of the human body [33, 34]. The possible reasons especially why the taxonomic composition of BV-associated bacteria is so different for each subject are: 1) differences in host immune response; 2) differences in geographic location; 3) expression of legends for bacterial attachment to epithelial cells; 4) the chemical and physical environments of the host; 5) intra- and inter specific microbial competition.

We identified several non-lactobacillus bacteria associated with BV, including *Atopobium, Prevotella, Gemella, Anaerococcus, Parvimonnas* and *Peptoniphilus,* and *Dialister, Sneathia*. At phylum lever, BV-group consists of mostly *Firmicutes* while BV+ group also has higher abundance in *Bacteroidetes* and *Actinobacteria*. We found that species diversity was more richness in BV+ group.

The preferential occurrence of several species in subjects of BV+ group was tested by t-test and analyzed strictly according to presence of absence of the species in the subject by chi-test. The results showed that most of these species were more prevalent in BV+ group. *Gardnerella vaginalis* was first found to be associated with BV. *Gardnerella vaginalis* was not more prevalent in BV+ group, but its presence was more common in BV+ group. Notably, Among the *Lactobacillus* species, Only *L. iners* and *L. coleohominis* were found with overall occurrence of 12.16% and 0.05% for BV + group. At genus level, the subjects in BV- groups have fewer bacterial genera (3.9 ± 3.35) than BV+ groups (11.8 ± 4.44). Several genera were found predominantly including *Atopobium, Porphyromonas, Prevotella, Parvimonas, Peptoniphilus, Dialister, Aerococcus, Acetivibrio BVAB3, Leptotrichia* and *Sneathia*

in BV+ group, this had been previously identified as BV-associated bacteria [23-26].

In previous reports about the microbial ecosystem in human gastrointestinal tract showed that the host immune system and the cell surface receptors may influence human–microbial symbioses and dictate adhesion of microorganisms [35,36]. McFall-Ngai [37] put forward the point of view that adaptive immunity plays a role in recognizing and managing the complex community composition of beneficial microorganisms in vertebrates. Based on the above, we proposed that similar selectivity may occur in vagina and the local vaginal immune system may play an important role in determining the composition of vaginal microbial communities.

Conclusions

The data presented here about the composition and richness of the vaginal microbial ecosystem in BV and health state will provide insight into the etiology of BV and the basis for the prevention and control of BV of microbial ecosystem in women.

Acknowledgements

This work was supported by the start grant from Guangdong province and Southern Medical University (to Shuhong Luo). We thank Yilong Luo, Yanghui Zheng and Yang Li for the sample collection. We thank Dr. Mengfei Ho and Dr. Brenda A, Wilson for critic reading of the manuscript.

Authors' contributions

SL and ML participated in the design of the study, the analysis of the data. YL, WH, CH, WX, SZ, JT and XL participated in the sample collection, DNA extraction, PCR library construction, sequencing and interpretation of the data. SL, ML and YL wrote the manuscript. All authors read and approved the final manuscript.

References

1. Sobel JD (2000) Bacterial vaginosis. Annu Rev Med 51: 349-356.
2. Haggerty CL, Hillier SL, Bass DC, Ness RB; PID Evaluation and Clinical Health study investigators (2004) Bacterial vaginosis and anaerobic bacteria are associated with endometritis. Clin Infect Dis 39: 990-995.
3. Taha TE, Hoover DR, Dallabetta GA, Kumwenda NI, Mtimavalye LA, et al. (1998) Bacterial vaginosis and disturbances of vaginal flora: association with increased acquisition of HIV. AIDS 12: 1699-1706.
4. Verstraelen H, Verhelst R, Claeys G, De Backer E, Temmerman M, et al. (2009) Longitudinal analysis of the vaginal microflora in pregnancy suggests that L. crispatus promotes the stability of the normal vaginal microflora and that L. gasseri and/or L. iners are more conducive to the occurrence of abnormal vaginal microflora. BMC Microbiol 9: 116.
5. Hyman RW, Fukushima M, Diamond L, Kumm J, Giudice LC, et al. (2005) Microbes on the human vaginal epithelium. Proc Natl Acad Sci U S A 102: 7952-7957.
6. Fredricks DN, Marrazzo JM (2005) Molecular methodology in determining vaginal flora in health and disease: its time has come. Curr Infect Dis Rep 7: 463-470.
7. Amsel R, Totten PA, Spiegel CA, Chen KC, Eschenbach D, et al. (1983) Nonspecific vaginitis. Diagnostic criteria and microbial and epidemiologic associations. Am J Med 74: 14-22.
8. Kim TK, Thomas SM, Ho M, Sharma S, Reich CI, et al. (2009) Heterogeneity of vaginal microbial communities within individuals. J Clin Microbiol 47: 1181-1189.
9. Lane DJ (1991) 16S/23S rRNA sequencing. in Nucleic Acid Techniques in Bacterial Systematics, eds Stackebrandt E, Goodfellow M (Wiley, New York), , pp 115–175.
10. Edwards U, Rogall T, Blöcker H, Emde M, Böttger EC (1989) Isolation and direct complete nucleotide determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. Nucleic Acids Res 17: 7843-7853.
11. Wilson KH, Blitchington RB, Greene RC (1990) Amplification of bacterial 16S ribosomal DNA with polymerase chain reaction. J Clin Microbiol 28: 1942-1946.
12. Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8: 175-185.
13. Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 8: 186-194.
14. Chou HH, Holmes MH (2001) DNA sequence quality trimming and vector removal. Bioinformatics 17: 1093-1104.
15. Cole JR, Wang Q, Cardenas E, Fish J, Chai B, et al. (2009) The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res 37: D141-145.
16. Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73: 5261-5267.
17. Field JG, Clarke KR, Warwick RM (1982) A Practical Strategy for Analyzing Multispecies Distribution Patterns. Mar Ecol Prog Ser 8: 37-52.
18. Bray JR, Curtis JT (1957) An Ordination of the Upland Forest Communities of Southern Wisconsin. Ecol Monog 27: 325-349.
19. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, et al. (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75: 7537-7541.
20. Roesch LF, Fulthorpe RR, Riva A, Casella G, Hadwin AK, et al. (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. ISME J 1: 283-290.
21. Bunge J, M Fitzpatrick (1993) Estimating the number of species: A review. Journal of the American Statistical Association, 88, 364-373.
22. Gardner HL, Dukes CD (1955) Haemophilus vaginalis vaginitis: a newly defined specific infection previously classified non-specific vaginitis. Am J Obstet Gynecol 69: 962-976.
23. Fredricks DN, Fiedler TL, Marrazzo JM (2005) Molecular identification of bacteria associated with bacterial vaginosis. N Engl J Med 353: 1899-1911.
24. Verhelst R, Verstraelen H, Claeys G, Verschraegen G, Delanghe J, et al. (2004) 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 4: 16.
25. Oakley BB, Fiedler TL, Marrazzo JM, Fredricks DN (2008) Diversity of human vaginal bacterial communities and associations with clinically defined bacterial vaginosis. Appl Environ Microbiol 74: 4898-4909.
26. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, et al. (2011) Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci U S A 108 Suppl 1: 4680-4687.
27. Shi Y, Chen L, Tong J, Xu C (2009) Preliminary characterization of vaginal microbiota in healthy Chinese women using cultivation-independent methods. J Obstet Gynaecol Res 35: 525-532.
28. Ling Z, Kong J, Liu F, Zhu H, Chen X, et al. (2010) Molecular analysis of the diversity of vaginal microbiota associated with bacterial vaginosis. BMC Genomics 11: 488.
29. Jan Leps, Petr Smilauer (2003) Using the Canoco for Windows 4.5 package. Multivariate Analysis of Ecological Data using CANOCO edited by Cambridge University Press, ,43-58
30. Zhou X, Bent SJ, Schneider MG, Davis CC, Islam MR, et al. (2004) Characterization of vaginal microbial communities in adult healthy women using cultivation-independent methods. Microbiology 150: 2565-2573.
31. Zhou X, Brown CJ, Abdo Z, Davis CC, Hansmann MA, et al. (2007) Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women. ISME J 1: 121-133.
32. Zhou X, Hansmann MA, Davis CC, Suzuki H, Brown CJ, et al. (2010) The vaginal bacterial communities of Japanese women resemble those of women in other racial groups. FEMS Immunol Med Microbiol 58: 169-181.
33. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, et al. (2005) Diversity of the human intestinal microbial flora. Science 308: 1635-1638.

34. Gao Z, Tseng CH, Pei Z, Blaser MJ (2007) Molecular analysis of human forearm superficial skin bacterial biota. Proc Natl Acad Sci U S A 104: 2927-2932.

35. Mazmanian SK, Liu CH, Tzianabos AO, Kasper DL (2005) An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. Cell 122: 107-118.

36. Dethlefsen L, McFall-Ngai M, Relman DA (2007) An ecological and evolutionary perspective on human-microbe mutualism and disease. Nature 449: 811-818.

37. McFall-Ngai M (2007) Adaptive immunity: care for the community. Nature 445: 153.