Characterization of vaginal microbiota in Thai women

Auttawit Sirichtoat 1, Pranom Buppasiri 2, Chulapan Engchanil 1, Wises Namwat 1, Kiatichai Faksri 1, Nipaporn Sankuntaw 3, Ekawat Pasomsub 4, Wasun Chantratita 5, Viraphong Lulitanond Corresponding Author

1 Department of Microbiology and Research and Diagnostic Center for Emerging Infectious Diseases, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand
2 Department of Obstetrics and Gynecology, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand
3 Chulabhorn International College of Medicine, Thammasat University, Pathum Thani, Thailand
4 Division of Virology, Department of Pathology, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand
5 Medical Genome Center, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand

Corresponding Author: Viraphong Lulitanond
Email address: viraphng@kku.ac.th

Background: The vaginal microbiota (VMB) plays a key role in women’s reproductive health. VMB composition varies with ethnicity, making it necessary to characterize the VMB of the target population before interventions to maintain and/or improve the vaginal health are undertaken. Information on the VMB of Thai women is currently unavailable. We therefore characterized the VMB in normal Thai women. Methods: Vaginal samples derived from 25 Thai women were subjected to 16S rRNA gene next-generation sequencing (NGS) on the Ion Torrent PGM platform. Results: Two groups of VMB were detected, lactobacilli-dominated (LD) and non-lactobacilli dominated (NLD) groups. Lactobacillus iners was the most common species found in the LD group while Gardnerella vaginalis followed by Atopobium vaginae and Pseudomonas stutzeri were commonly found in the NLD group. Conclusions: The VMB patterns present in normal Thai women is essential information to further determine the factors associated with VMB patterns in vaginal health and disease and to develop proper management of reproductive health of Thai women.
Characterization of vaginal microbiota in Thai women

Auttawit Sirichoat¹, Pranom Buppasiri², Chulapan Engchanil¹, Wises Namwat¹, Kiatichai Faksri¹, Nipaporn Sankuntaw³, Ekawat Pasomsub⁴, Wasun Chantratita⁵, Viraphong Lulitanond¹

¹ Department of Microbiology and Research and Diagnostic Center for Emerging Infectious Diseases, Faculty of Medicine, Khon Kaen University, Khon Kaen, 40002, Thailand.
² Department of Obstetrics and Gynecology, Faculty of Medicine, Khon Kaen University, Khon Kaen, 40002, Thailand.
³ Chulabhorn International College of Medicine, Thammasat University, Pathum Thani, 12120, Thailand.
⁴ Division of Virology, Department of Pathology, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, 10400, Thailand.
⁵ Medical Genome Center, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, 10400, Thailand.

Corresponding author:
Viraphong Lulitanond¹
Department of Microbiology, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand
40002
E-mail: viraphng@kku.ac.th; viraphng@gmail.com
Abstract

Background: The vaginal microbiota (VMB) plays a key role in women’s reproductive health. VMB composition varies with ethnicity, making it necessary to characterize the VMB of the target population before interventions to maintain and/or improve the vaginal health are undertaken. Information on the VMB of Thai women is currently unavailable. We therefore characterized the VMB in normal Thai women.

Methods: Vaginal samples derived from 25 Thai women were subjected to 16S rRNA gene next-generation sequencing (NGS) on the Ion Torrent PGM platform.

Results: Two groups of VMB were detected, lactobacilli-dominated (LD) and non-lactobacilli dominated (NLD) groups. Lactobacillus iners was the most common species found in the LD group while Gardnerella vaginalis followed by Atopobium vaginae and Pseudomonas stutzeri were commonly found in the NLD group.

Conclusions: The VMB patterns present in normal Thai women is essential information to further determine the factors associated with VMB patterns in vaginal health and disease and to develop proper management of reproductive health of Thai women.
Introduction

The vaginal microbiota (VMB) plays an important role in vaginal health and in particular it has a protective role against vaginal infection. The composition of the VMB varies greatly among ethnic groups (Ravel et al., 2011). Up to now there is no definite study about how the VMB varies by ethnicity, however, some studies suggested the differences in VMB by ethnicities may be related to host genetic variations, which associated with individual differences in immune response (Blekhman et al., 2015; Green et al., 2015; Ma et al., 2014). In addition, VMB composition can be affected by various internal and external factors, e.g. host physiology, transitional reproductive period, menstruation, pregnancy, infections, birth control procedures and amount of sexual activity (Eschenbach et al., 2000; Gajer et al., 2012; Huang et al., 2014; Romero et al., 2014; Smith & Ravel, 2017; Vodstrcil et al., 2017). In reproductive-aged women, the VMB is dominated by various Lactobacillus species such as L. iners, L. crispatus, L. gasseri, and L. jensenii (Gajer et al., 2012; Ravel et al., 2011). In vitro studies found that Lactobacilli can inhibit pathogens through the production of hydrogen peroxide and lactic acid which could maintain an acidic environment in the vagina (Aroutcheva et al., 2001; Eschenbach et al., 1989).

Several studies have characterized the VMB in healthy, reproductive women of different ethnicities using DNA sequence-based methods, especially next-generation sequencing (NGS) techniques (Borgdorff et al., 2017; Fettweis et al., 2014; Ravel et al., 2011). Five community state types (CSTs) have been recognized based on bacterial relative abundance on the basis of 16S ribosomal RNA (rRNA) sequences obtained using NGS (Ravel et al., 2011; Zhou et al., 2004). Four of these CSTs, i.e. CST-I, CST-II, CST-III and CST-V, were dominated by L. crispatus, L. gasseri, L. iners and L. jensenii, respectively. An additional type, CST-IV, is comprised of a high diversity of strict and facultative anaerobic bacteria including members of
the genera *Gardnerella, Atopobium, Corynebacterium, Prevotella, Mobiluncus, Anaerococcus, Sneathia* and others, which are associated with vaginal infections especially bacterial vaginosis (BV) ([Fredricks et al., 2005; Ling et al., 2010; Ravel et al., 2011; Swidsinski et al., 2005]). NGS-based studies have previously explored the VMB in Asian, African and European women ([Borgdorff et al., 2017; Chaban et al., 2014; Fettweis et al., 2014; Ravel et al., 2011; van de Wijgert et al., 2014]). Among Asian populations, VMB in Chinese ([Ling et al., 2013; Shi et al., 2009]), Japanese ([Yoshimura et al., 2011; Zhou et al., 2010]) and South Korean ([Hong et al., 2016; Lee et al., 2013]) populations have already been studied, but women in South East Asia, including Thailand, have not been investigated. To our knowledge, this is the first report to characterize the VMB in healthy Thai women based on 16S rRNA gene sequences obtained using NGS.

**Materials and methods**

**Subject selection and sample collection**

Twenty-five women attending the gynecologic out-patient clinics for cervical cancer screening, Srinagarind Hospital, Faculty of Medicine, Khon Kaen University, were enrolled in this study. The inclusion criteria used for subject selection specified non-pregnant women aged 20-45 years with regular menstruation, with normal vaginal examination and without any clinical symptoms on examination by one gynecologist. The exclusion criteria were women with the following conditions: having menstruation or sexual intercourse within the previous 24 hours, use of antibiotics or vaginal antimicrobials within the last month, presence of any intravaginal product, use of a vaginal douche in the past week, and active infection or diagnosis of any defined sexually transmitted disease. The subjects selected by the above criteria were defined as
normal women. All participants provided written informed consent and filled out a questionnaire
detailing their level of education, employment, sociodemographic characteristics, dietary habits,
reproductive health habits, and sexual behaviors (including number of children), income per
month, alcohol consumption per week, exercise per week, sexual activity per week, number of
sexual partners in the past, regularity of their menstrual cycles in the last six months, current
hormone contraceptive use, douching practice in the past month, abnormal vaginal discharge in
the past month and female sterilization. This study was approved by the Ethics Committee of
Khon Kaen University (HE601017).

A total of four points at mid right and left, anterior and posterior vaginal wall were
swabbed with sterile cotton-tipped applicators by one gynecologist. The swabs were placed in
sterile containers with phosphate-buffer saline (PBS) and stored at -80°C until analysis.

Genomic DNA extraction from vaginal swabs

Genomic DNA was extracted using the method described by Ravel et al. (2011) with
slight modification. In briefly, the vaginal swabs were thawed on ice and adhered material
resuspended by vigorous vortexing for 5 min. Bacterial cells were lysed by incubation for 1 hour
at 37°C in TE50 buffer (10 mM Tris HCl, 50 mM EDTA, pH 8.0) containing 10 mg/mL of
lysozyme (Merck, Darmstadt, Germany), 25,000 U/mL of mutanolysin (Sigma Aldrich, USA, St.
Louis) and 4,000 U/mL of lysostaphin (Sigma-Aldrich), followed by beating with 40-400 µM
glass beads (NucleoSpin® Bead Tubes Type B, Macherey-Nagel, Düren, Germany) at speed 10
for 2 min with air cooling using the Bullet Blender® Blue Homogenizer (BBX 24B, Next
Advance, Inc., USA). After bead-beating, the crude lysates were processed for purification of
genomic DNA using QIAamp® DNA Mini Kit (Qiagen, Hilden, Germany) according to the
manufacturer’s recommendations. Integrity, size and concentration of the purified genomic DNA was determined using Fragment Analyzer™ (Advanced Analytical Technologies, Inc., USA).

16S rRNA gene sequencing by NGS

16S rRNA gene sequencing was done as previously described (Barb et al., 2016) with minor modifications. The seven 16S hypervariable regions were amplified with 2 primer sets, one set for amplifying the V2, V4, V8 hypervariable regions and the other set for amplifying the V3, V6-7, V9 hypervariable regions. Primers were supplied with the Ion 16S™ Metagenomics Kit (Life Technologies, Grand Island, NY) and used according to the manufacturer’s recommendations. The combined PCR products were processed to make the DNA library by using the Ion Plus™ Fragment Library Kit (Life Technologies) with sample indexing using the Ion Xpress™ Barcodes Adapters 1-32 Kit (Life Technologies). The adapter-ligated and nick-repaired DNA was amplified with the following steps: 1 hold at 95°C for 5 min; 5-7 cycles of 95°C for 15 sec, 58°C for 15 sec, 70°C for 1 min; hold at 4°C. The PCR products were then purified using Agencourt® AMPure® XP Reagent (Beckman Coulter, Inc, Atlanta, Georgia). The processed libraries were quantified using a Bioanalyzer® instrument (Agilent Technologies, Santa Clara, CA.) with the use of the Agilent® High-Sensitivity DNA Kit. Equal volumes of all 25 samples were pooled and processed with the Ion PGM™ Hi-Q™ OT2 Kit in the Ion OneTouch™ 2 System (Life Technologies) according to the manufacturer’s instructions.

Sequencing was performed in an Ion 318™ Chip (Life Technologies) using the Ion PGM™ Sequencing 400 Kit on the Ion PGM™ System (Life Technologies).

Sequence data analysis
Base calling and run demultiplexing were performed using Torrent Suite™ Software version 5.0.5 (Life Technologies) with default parameters. All sequences were trimmed according to the quality. To pass, a sequence read had the following criteria: 1) a perfect match to a barcode sequence and the primer; 2) was at least 200 bp in length; 3) had no more than two undetermined bases; and 4) had an average quality score greater than Q20. Sequences were analyzed using Ion Reporter™ Software version 5.6 with the 16S Metagenomics workflow module (Life Technologies). Clustering of operational taxonomic units (OTUs) and taxonomic classification were performed based on two-step Basic Local Alignment Search Tool (BLAST) that maps to two separate reference libraries. In the first step, reads were aligned to the MicroSEQ® 16S Reference Library v2013.1 database (Woo et al., 2003). Subsequently, any unaligned reads subject to second alignment to the Greengenes v13.5 database (DeSantis et al., 2006). At least 10 of unique reads were valid and ≥90% for the alignment coverage needed between hit and query. Genus- and species-level identifications were accepted at ≥97% and ≥99% sequence identity, respectively (Drancourt et al., 2000). In case of any unusual microorganism, the sequence reads were reanalyzed by BLAST with the National Center for Biotechnology Information (NCBI) taxonomy database. Any OTUs represented by only one sequence (singletons) and those with fewer than 10 reads in the sample were excluded from further analysis. Alpha- (within sample) and beta- (among samples) diversities were calculated using Quantitative Insights Into Microbial Ecology (QIIME) software (Caporaso et al., 2010). Alpha diversity metrics, consisting of the Shannon diversity index, observed species and Chao1 index, were graphed using GraphPad Prism version 5.01 (GraphPad Software Inc., USA). The beta diversity metric was calculated according to Bray-Curtis dissimilarity index (Bray & Curtis, 1957) and displayed through a principal coordinates analysis (PCoA) plot. Heat map and
hierarchical clustering tree were generated using R Studio (Version 1.1.414, Boston, USA) employing R version 3.4.1 (R Development Core Team). Raw sequences were deposited in the NCBI Sequence Read Archive (SRA) accession number SRP158176.

Statistical analysis

Statistical analysis was performed using IBM SPSS Statistics for Windows, version 19.0 software (SPSS Inc., Chicago, IL, USA). Demographic characteristics between two groups were compared using the chi-squared or Fisher’s exact test. Mann-Whitney U test for non-parametric data was used for bacterial abundance and diversity comparisons between two groups. \( P \)-values less than 0.05 were considered as statistically significant.

Results

Participant characteristics

The median age and body mass index of the enrolled subjects were 39 years (ranging from 31 to 45) and 21.2 kg/m\(^2\) (ranging from 16.0 to 35.8), respectively. Subjects fell into two groups: lactobacilli-dominated (LD) and non-lactobacilli dominated (NLD) groups (see below). No significant association between sociodemographic characteristics and VMB group (LD or NLD) was found, as shown in Table 1.

Analysis of VMB in the studied subjects

High-throughput sequencing data of all vaginal samples derived from Ion Torrent PGM yielded a total of 4,775,332 raw sequences with an average of 191,013 reads per sample (ranging from 33,969 to 382,574 reads). We discarded any reads shorter than 150 bp and those OTUs represented fewer than 10 times in the sample. This left a total of 4,189,265 valid reads with an
average of 167,570 reads per sample (ranging from 28,825 to 333,480 reads) for analysis. The average length of included reads was 243 bp (ranging from 237 to 253 bp). A total of 3,553,612 mapped reads were clustered into OTUs and given taxonomic identification. Forty-one genera and 72 species were identified (Table S1 and S2). VMB were clustered into two distinct groups: LD (n=14) and NLD (n=11). In the LD group, the average of relative abundance of *Lactobacillus* species is 90% and the relative abundance of this species among the LD individuals is ≥60% and could be further divided into two subgroups: one dominated by *L. iners* (n=10) and the second subgroup (n=4) was dominated by *L. crispatus, L. gasseri, L. jensenii* and *L. johnsonii*. The NLD group could similarly be divided into two subgroups: one dominated by *Gardnerella vaginalis* (n=3) and the other (n=8) dominated by anaerobic bacteria of a mixture of several species including *G. vaginalis, Atopobium vaginae* and *Pseudomonas stutzeri*. Fig. 1 represents the hierarchical clustering tree (Fig. 1A) and the heat map (Fig. 1B) of bacterial taxa identified from 25 vaginal samples. Table S3 indicates the mean percentage representation of each bacterial genus present in these two groups.

In this study, alpha-diversity for determination of bacterial diversity within individual women used three metrics: Shannon diversity index (estimated evenness and richness), observed species (observed bacterial richness) and Chao1 (estimated bacterial richness). The Shannon diversity index differed significantly between the two groups (P<0.001) (Fig. 2). The observed and estimated numbers of OTUs in individuals in the NLD group were significantly higher than those in the LD group (both P<0.001) (Fig. 3). All these results indicated that the bacterial communities in NLD individuals have more species diversity than those in LD individuals. In beta-diversity analysis, the PCoA plot based on taxonomic profiles showed that some distinct clustering appeared to overlap between two groups (Fig. 4). This result suggested that the VMB
of both LD and NLD individuals did not tightly cluster by group. In addition, the relative
distances based on the Bray-Curtis dissimilarity index showed significantly greater distances
among LD individual samples than among NLD individual samples (0.82±0.10 versus
0.72±0.09, P=0.021). Since the distance measurement indicates the degree of similarity among
samples (van de Wijgert & Jespers, 2016), this result indicates that VMB in LD individuals
resemble one another more closely than do those in NLD individuals.

The number of species unique to either LD or NLD, or common to both groups, were 14
(19.44%), 41 (56.95%) and 17 (23.61%), respectively (Fig. 5). The 17 common bacterial species
found in both groups were A. vaginae, Finegoldia magna, G. vaginalis, L. crispatus, L. gasseri,
L. iners, L. johnsonii, L. kefiranofaciens, L. vaginalis, Megasphaera spp., Peptoniphilus harei,
Peptostreptococcus anaerobius, Prevotella bivia, P. timonensis, Pseudomonas stutzeri,
Streptococcus intermedius and Ureaplasma parvum. Fig. 6 represents the most abundant
bacterial species detected in the vaginal samples of LD and NLD groups. G. vaginalis, A.
vaginae and P. bivia were significantly associated with the NLD group (P<0.05) while L. iners
was significantly associated with the LD group (P<0.05). The bacterial species with abundance
less than 1% are F. magna, L. kefiranofaciens, L. vaginalis, Peptoniphilus harei,
Peptostreptococcus anaerobius, S. intermedium and U. parvum.

Discussion

The community composition of the VMB can vary greatly among women. However,
although it can be absent entirely, Lactobacillus is likely to be the dominant, and sometimes
only, genus in VMB (Cherpes et al., 2008b; Linhares et al., 2011; Ma et al., 2012; Petrova et
al., 2015; Smith & Ravel, 2017). We found the VMB of normal Thai women to fall into two
groups. In the LD group, lactobacilli, and especially \textit{L. iners}, can make up 90\% of the bacterial cells present. In the NLD group, \textit{G. vaginalis} followed by \textit{A. vaginae} and \textit{Pseudomonas stutzeri} jointly constitute the majority of bacteria present. Previous studies found the difference on the composition of VMB among reproductive age women of different ethnicity (\textit{Fettweis et al., 2014; Ravel et al., 2011}). Studied on VMB of asymptomatic North America women by \textit{Ravel et al. (2011)} using Roche 454 sequencing (454 sequencing) of the V1-V2 regions of 16S rRNA gene found that \textit{Lactobacillus}-dominated (80-90\%) is the most common type of VMB in Asian and White women but less common (approx.. 60\%) in Hispanic and African American women. Furthermore, a \textit{Lactobacillus}-dominated VMB was also found in normal Chinese and Korean women by using 454 sequencing of the V3 and V3-V5 regions, respectively (\textit{Hong et al., 2016; Ling et al., 2010}). A high abundance of \textit{Lactobacillus} species is generally considered as a good biomarker for a healthy vaginal ecosystem (\textit{Petrova et al., 2015}). Lactobacilli promote a protective environment in the vagina by several mechanisms, such as production of hydrogen peroxide, bacteriocins or certain metabolites, that can inhibit the colonization and growth of various vaginal pathogens (\textit{Aroutcheva et al., 2001; Eschenbach et al., 1989}).

\textit{Lactobacillus iners}, the dominant species of VMB in our study, is also the most common species in asymptomatic North American women with Asian ethnicity and in African American women by using 454 sequencing of the V1-V2 and V1-V3 regions of 16S rRNA genes, respectively (\textit{Fettweis et al., 2014; Ravel et al., 2011}). Previous studies with different sequencing platforms of 16S rRNA gene, e.g. Illumina (V3-V4) and 454 (V1-V2 or V1-V3) sequencing, found other \textit{Lactobacillus} species such as \textit{L. crispatus}, \textit{L. gasseri} and \textit{L. jensenii} are commonly the dominant species in White or European women, but are less common in Asian, Black and Hispanic women (\textit{Borgdorff et al., 2017; Fettweis et al., 2014; Ravel et al., 2011}). The
difference in VMB composition likely have a genetic and ethnic basis (Green et al., 2015).

Nevertheless, a systematic review by van de Wijgert et al. (2014) found that L. iners is present in most women, either healthy or with BV, while L. crispatus is commonly found only in healthy women. Longitudinal studies have found that women who have L. iners-dominated VMB at baseline are more likely to transition to a BV-associated VMB than women who have L. crispatus-dominated VMB. Lactobacillus crispatus-dominated VMB more often transitions to an L. iners-dominated or mixed lactobacilli VMB than to a BV-associated VMB (Gajer et al., 2012). The role of L. iners remains controversial. It appears that L. iners is well adapted to the vaginal niche, able to survive in a BV-like environment and often persisting after antibiotic treatment (Huang et al., 2014; van de Wijgert et al., 2014). It may help to restore a lactobacilli-dominated microbiota during and after dysbiosis and/or after antibiotic treatment (van de Wijgert et al., 2014). Thus, L. iners-dominated VMB might indicate a transitional condition between normal and abnormal. However, since our participants with L. iners-dominated VMB show no symptom both before and after sample collection perhaps this is “normal” for Thai women.

Other studies have proposed that genetic/ethnic differences in immune responses might make the vaginal mucosa more favorable for colonization by L. iners over L. crispatus in African women (Doh et al., 2004; Nguyen et al., 2004; Ryckman et al., 2008). However, the exact mechanism leading to transitions in the microbiota are not known and might be affected by various factors, e.g. the hormonal changes, sexual intercourse, smoking, personal hygiene, antibiotic treatment, ethnicity, individual differences in immune status/response, vaginal epithelial cells and their secretions (Petrova et al., 2015; Petrova et al., 2017; van de Wijgert et al., 2014).

Nearly half of the subjects (11 of the 25 subjects) was dominated by a suite of species, not by a single taxon. This type of VMB was previously found in 40.4%, 19.8% and 38.5% of
asymptomatic Black or African women (Ravel et al., 2011), Asian women (Ravel et al., 2011), and women in Amsterdam (Borgdorff et al., 2017), respectively. Previous systematic reviews have revealed that some healthy women have this type of VMB (Huang et al., 2014; Petrova et al., 2015), which is dominated by members of the genera Gardnerella, Atopobium, Prevotella, Corynebacterium, Anaerococcus, Peptoniphilus, Mobiluncus and Sneathia. These species have evolved mechanisms to persist in a slightly alkaline environment and to adhere to the vaginal epithelial cells. It has been suggested that such a suite of bacteria is able to maintain the protective function of the vagina through lactic acid production to lower vaginal pH (Gajer et al., 2012). Lactic acid can be produced by either homolactic or heterolactic acid fermentation by certain facultatively or strictly anaerobic bacteria such as Atopobium, Streptococcus, Staphylococcus, Megashara and Leptotrichia (Petrova et al., 2015; Zhou et al., 2004). These bacteria might contribute to maintain the balance of the healthy vaginal ecosystem. Thus, the polymicrobial VMB seen in our and other studies indicated that this type of VMB can be present in normal healthy women (Petrova et al., 2015).

It should be noted that in the study of VMB using 16S rRNA gene sequencing there are difference in the usage of 16S rRNA region to generate reads and in the database for taxonomic classification among the studies. There is much speculation on which hypervariable region provides sufficient sequences diversity to identify the most bacteria accurately (Liu et al., 2008; Schloss et al., 2011). Many studies found that using different hypervariable regions of the gene provides variable results (Barb et al., 2016; Chakravorty et al., 2007; Youssef et al., 2009). The best solution would be to sequence the entire 16S rRNA gene; however, this approach is not possible with short read NGS platforms. In this study, we used Ion Torrent PGM platform which using multiple variable regions (V2-V4 and V6-V9 regions) to explore the VMB which would
provide sufficient sequence to identify the most bacteria accurately. In addition, in the taxonomic
classification in this study primarily performed by alignment to the MicroSEQ® 16S Reference
Library v2013.1 database and the Greengenes v13.5 database as described in Materials and
Methods above revealed some unusual microorganisms, i.e. *Olsenella umbonata* and *O. profusa*.
The sequence reads of these two microorganisms were reanalyzed with NCBI database and
found them to belong to the most common species, i.e. *A. vaginae* (% identity of *O. umbonata*
and *A. vaginae* are 99 and 100, respectively, and the identity of both *O. profusa* and *A. vaginae* is
99%). For the other microorganisms, the classification using either MicroSEQ® 16S Reference
Library/Greengenes database or NCBI database found to give similar results.

Previous studies suggested that differences in bacterial composition in VMB might be
associated with both intrinsic and extrinsic factors, e.g. ethnicity, sociodemography,
environment, pregnancy, smoking status, sexual behavior, number of sexual partners, alcohol
consumption and host genetic factors (*Borgdorff et al., 2017; Fettweis et al., 2014; Huang et al.,
2014*). Five different VMB clusters were described among four different ethnic groups (White,
Black, Hispanic and Asian women) in North America (*Ravel et al., 2011*). Two additional
clusters, dominated by *G. vaginalis* or BV-associated bacteria (BVAB) type 1, were found in
another study in African American women and women of European ancestry (*Fettweis et al.,
2014*). Genomic markers have revealed that differences in VMB by ethnicity might be related to
mitochondrial DNA (mtDNA) polymorphisms (*Blekhman et al., 2015; Green et al., 2015; Ma et
al., 2014*). High estradiol during pregnancy has been reported to be associated with high levels of
lactobacilli and low bacterial diversity (*Aagaard et al., 2012; Petricevic et al., 2012*), and with
low incidence of BV and low prevalence of BVAB during pregnancy (*Romero et al., 2014*).
Smoking is a risk factor of BV (*Cherpes et al., 2008a; Ryckman et al., 2009*) and smokers appear
to have a lower proportion of vaginal *Lactobacillus* spp. than do non-smokers (*Brotman et al., 2014*). Sexual behaviors and number of sexual partners have been implicated the bacterial composition and diversity of the VMB (*Huang et al., 2014; van de Wijgert et al., 2014*). Prior studies have shown that the bacterial diversity in women engaging in high-risk sexual behavior was increased along with the decline of lactobacilli (*Wessels et al., 2017*). Lastly, due to a small sample size in this study, the large sample size in future study is inevitably required to draw a more meaningful conclusion.

**Conclusion**

Our study is the first to report the types of VMB in normal Thai women using 16S rRNA gene sequence data obtained by NGS. Two major groups were recognized, i.e. LD and NLD groups. *Lactobacillus iners* is the dominant species of VMB in the LD group while several species are abundant in VMB in the NLD group. The information on VMB in Thai women is a starting point for further studying of factors involved in the development and maintenance of VMB communities in vaginal health and disease. A better understanding of VMB, including their interaction with external and internal factors, will assist in development of effective strategies to manage reproductive health of Thai women.

**Acknowledgements**

We would like to acknowledge Professor David Blair for editing the manuscript via Publication Clinic KKU, Thailand.
References

Aagaard K, Riehle K, Ma J, Segata N, Mistretta TA, Coarfa C, Raza S, Rosenbaum S, Van den Veyver I, Milosavljevic A, Gevers D, Huttenhower C, Petrosino J, and Versalovic J. 2012. A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. *PLoS One* 7:e36466 DOI: 10.1371/journal.pone.0036466.

Aroutcheva A, Gariti D, Simon M, Shott S, Faro J, Simoes JA, Gurguis A, and Faro S. 2001. Defense factors of vaginal lactobacilli. *American Journal of Obstetrics and Gynecology* 185:375-379 DOI: 10.1067/mob.2001.115867.

Barb JJ, Oler AJ, Kim HS, Chalmers N, Wallen GR, Cashion A, Munson PJ, and Ames NJ. 2016. Development of an Analysis Pipeline Characterizing Multiple Hypervariable Regions of 16S rRNA Using Mock Samples. *PLoS One* 11:e0148047 DOI: 10.1371/journal.pone.0148047.

Blekhman R, Goodrich JK, Huang K, Sun Q, Bukowski R, Bell JT, Spector TD, Keinan A, Ley RE, Gevers D, and Clark AG. 2015. Host genetic variation impacts microbiome composition across human body sites. *Genome Biology* 16:191 DOI: 10.1186/s13059-015-0759-1.

Borgdorff H, van der Veer C, van Houdt R, Alberts CJ, de Vries HJ, Bruisten SM, Snijder MB, Prins M, Geerlings SE, Schim van der Loeff MF, and van de Wijgert J. 2017. The association between ethnicity and vaginal microbiota composition in Amsterdam, the Netherlands. *PLoS One* 12:e0181135 DOI: 10.1371/journal.pone.0181135.

Bray JR, and Curtis JT. 1957. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs* 27:325-349 DOI: 10.2307/1942268.
Brotman RM, He X, Gajer P, Fadrosh D, Sharma E, Mongodin EF, Ravel J, Glover ED, and Rath JM. 2014. Association between cigarette smoking and the vaginal microbiota: a pilot study. *BMC Infectious Diseases* 14:471 DOI: 10.1186/1471-2334-14-471.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttenhower C, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung MC, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, and Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7:335-336 DOI: 10.1038/nmeth.f.303.

Chaban B, Links MG, Jayaprakash TP, Wagner EC, Bourque DK, Lohn Z, Albert AY, van Schalkwyk J, Reid G, Hemmingsen SM, Hill JE, and Money DM. 2014. Characterization of the vaginal microbiota of healthy Canadian women through the menstrual cycle. *Microbiome* 2:23 DOI: 10.1186/2049-2618-2-23.

Chakravorty S, Helb D, Burday M, Connell N, and Alland D. 2007. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *Journal of Microbiological Methods* 69:330-339 DOI: 10.1016/j.mimet.2007.02.005.

Cherpes TL, Hillier SL, Meyn LA, Busch JL, and Krohn MA. 2008a. A delicate balance: risk factors for acquisition of bacterial vaginosis include sexual activity, absence of hydrogen peroxide-producing lactobacilli, black race, and positive herpes simplex virus type 2 serology. *Sexually Transmitted Diseases* 35:78-83 DOI: 10.1097/OLQ.0b013e318156a5d0.
Cherpes TL, Marrazzo JM, Cosentino LA, Meyn LA, Murray PJ, and Hillier SL. 2008b. Hormonal contraceptive use modulates the local inflammatory response to bacterial vaginosis. *Sexually Transmitted Infections* 84:57-61 DOI: 10.1136/sti.2007.026625.

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, and Andersen GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology* 72:5069-5072 DOI: 10.1128/AEM.03006-05.

Doh K, Barton PT, Korneeva I, Perni SC, Bongiovanni AM, Tuttle SL, Skupski DW, and Witkin SS. 2004. Differential vaginal expression of interleukin-1 system cytokines in the presence of *Mycoplasma hominis* and *Ureaplasma urealyticum* in pregnant women. *Infectious Diseases in Obstetrics and Gynecology* 12:79-85 DOI: 10.1080/10647440400003667.

Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral JP, and Raoult D. 2000. 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. *Journal of Clinical Microbiology* 38:3623-3630.

Eschenbach DA, Davick PR, Williams BL, Klebanoff SJ, Young-Smith K, Critchlow CM, and Holmes KK. 1989. Prevalence of hydrogen peroxide-producing *Lactobacillus* species in normal women and women with bacterial vaginosis. *Journal of Clinical Microbiology* 27:251-256.

Eschenbach DA, Thwin SS, Patton DL, Hooton TM, Stapleton AE, Agnew K, Winter C, Meier A, and Stamm WE. 2000. Influence of the normal menstrual cycle on vaginal tissue, discharge, and microflora. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 30:901-907 DOI: 10.1086/313818.
Fettweis JM, Brooks JP, Serrano MG, Sheth NU, Girerd PH, Edwards DJ, Strauss JF, 3rd, Vaginal Microbiome C, Jefferson KK, and Buck GA. 2014. Differences in vaginal microbiome in African American women versus women of European ancestry. *Microbiology* 160:2272-2282 DOI: 10.1099/mic.0.081034-0.

Fredricks DN, Fiedler TL, and Marrazzo JM. 2005. Molecular identification of bacteria associated with bacterial vaginosis. *The New England journal of medicine* 353:1899-1911 DOI: 10.1056/NEJMoa043802.

Gajer P, Brotman RM, Bai G, Sakamoto J, Schutte UM, Zhong X, Koenig SS, Fu L, Ma ZS, Zhou X, Abdo Z, Forney LJ, and Ravel J. 2012. Temporal dynamics of the human vaginal microbiota. *Science Translational Medicine* 4:132ra152 DOI: 10.1126/scitranslmed.3003605.

Green KA, Zarek SM, and Catherino WH. 2015. Gynecologic health and disease in relation to the microbiome of the female reproductive tract. *Fertility and Sterility* 104:1351-1357 DOI: 10.1016/j.fertnstert.2015.10.010.

Hong KH, Hong SK, Cho SI, Ra E, Han KH, Kang SB, Kim EC, Park SS, and Seong MW. 2016. Analysis of the Vaginal Microbiome by Next-Generation Sequencing and Evaluation of its Performance as a Clinical Diagnostic Tool in Vaginitis. *Annals of Laboratory Medicine* 36:441-449 DOI: 10.3343/alm.2016.36.5.441.

Huang B, Fettweis JM, Brooks JP, Jefferson KK, and Buck GA. 2014. The changing landscape of the vaginal microbiome. *Clinics in Laboratory Medicine* 34:747-761 DOI: 10.1016/j.cll.2014.08.006.
Lee JE, Lee S, Lee H, Song YM, Lee K, Han MJ, Sung J, and Ko G. 2013. Association of the vaginal microbiota with human papillomavirus infection in a Korean twin cohort. *PLoS One* 8:e63514 DOI: 10.1371/journal.pone.0063514.

Ling Z, Kong J, Liu F, Zhu H, Chen X, Wang Y, Li L, Nelson KE, Xia Y, and Xiang C. 2010. Molecular analysis of the diversity of vaginal microbiota associated with bacterial vaginosis. *BMC Genomics* 11:488 DOI: 10.1186/1471-2164-11-488.

Ling Z, Liu X, Luo Y, Wu X, Yuan L, Tong X, Li L, and Xiang C. 2013. Associations between vaginal pathogenic community and bacterial vaginosis in Chinese reproductive-age women. *PLoS One* 8:e76589 DOI: 10.1371/journal.pone.0076589.

Linhares IM, Summers PR, Larsen B, Giraldo PC, and Witkin SS. 2011. Contemporary perspectives on vaginal pH and lactobacilli. *American Journal of Obstetrics and Gynecology* 204:120 e121-125 DOI: 10.1016/j.ajog.2010.07.010.

Liu Z, DeSantis TZ, Andersen GL, and Knight R. 2008. Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. *Nucleic Acids Research* 36:e120 DOI: 10.1093/nar/gkn491.

Ma B, Forney LJ, and Ravel J. 2012. Vaginal microbiome: rethinking health and disease. *Annual Review of Microbiology* 66:371-389 DOI: 10.1146/annurev-micro-092611-150157.

Ma J, Coarfa C, Qin X, Bonnen PE, Milosavljevic A, Versalovic J, and Aagaard K. 2014. mtDNA haplogroup and single nucleotide polymorphisms structure human microbiome communities. *BMC Genomics* 15:257 DOI: 10.1186/1471-2164-15-257.

Nguyen DP, Genc M, Vardhana S, Babula O, Onderdonk A, and Witkin SS. 2004. Ethnic differences of polymorphisms in cytokine and innate immune system genes in pregnant
women. Obstetrics and Gynecology 104:293-300 DOI: 10.1097/01.AOG.0000133486.85400.5e.

Petricevic L, Domig KJ, Nierscher FJ, Krondorfer I, Janitschek C, Kneifel W, and Kiss H. 2012. Characterisation of the oral, vaginal and rectal Lactobacillus flora in healthy pregnant and postmenopausal women. European Journal of Obstetrics, Gynecology, and Reproductive Biology 160:93-99 DOI: 10.1016/j.ejogrb.2011.10.002.

Petrova MI, Lievens E, Malik S, Imholz N, and Lebeer S. 2015. Lactobacillus species as biomarkers and agents that can promote various aspects of vaginal health. Frontiers in Physiology 6:81 DOI: 10.3389/fphys.2015.00081.

Petrova MI, Reid G, Vaneechoutte M, and Lebeer S. 2017. Lactobacillus iners: Friend or Foe? Trends in Microbiology 25:182-191 DOI: 10.1016/j.tim.2016.11.007.

Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, Brotman RM, Davis CC, Ault K, Peralta L, and Forney LJ. 2011. Vaginal microbiome of reproductive-age women. Proceedings of the National Academy of Sciences of the United States of America 108 Suppl 1:4680-4687 DOI: 10.1073/pnas.1002611107.

Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Nikita L, Galuppi M, Lamont RF, Chaemsaiithong P, Miranda J, Chaiworapongsa T, and Ravel J. 2014. The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. Microbiome 2:4 DOI: 10.1186/2049-2618-2-4.

Ryckman KK, Simhan HN, Krohn MA, and Williams SM. 2009. Predicting risk of bacterial vaginosis: the role of race, smoking and corticotropin-releasing hormone-related genes. Molecular Human Reproduction 15:131-137 DOI: 10.1093/molehr/gan081.
Ryckman KK, Williams SM, Krohn MA, and Simhan HN. 2008. Racial differences in cervical cytokine concentrations between pregnant women with and without bacterial vaginosis. *Journal of Reproductive Immunology* 78:166-171 DOI: 10.1016/j.jri.2008.01.003.

Schloss PD, Gevers D, and Westcott SL. 2011. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS One* 6:e27310 DOI: 10.1371/journal.pone.0027310.

Shi Y, Chen L, Tong J, and Xu C. 2009. Preliminary characterization of vaginal microbiota in healthy Chinese women using cultivation-independent methods. *The Journal of Obstetrics and Gynaecology Research* 35:525-532 DOI: 10.1111/j.1447-0756.2008.00971.x.

Smith SB, and Ravel J. 2017. The vaginal microbiota, host defence and reproductive physiology. *The Journal of Physiology* 595:451-463 DOI: 10.1113/JP271694.

Swidsinski A, Mendling W, Loening-Baucke V, Ladhoff A, Swidsinski S, Hale LP, and Lochs H. 2005. Adherent biofilms in bacterial vaginosis. *Obstetrics and Gynecology* 106:1013-1023 DOI: 10.1097/01.AOG.0000183594.45524.d2.

van de Wijgert JH, Borgdorff H, Verhelst R, Crucitti T, Francis S, Verstraelen H, and Jespers V. 2014. The vaginal microbiota: what have we learned after a decade of molecular characterization? *PLoS One* 9:e105998 DOI: 10.1371/journal.pone.0105998.

van de Wijgert JH, and Jespers V. 2016. Incorporating microbiota data into epidemiologic models: examples from vaginal microbiota research. *Annals of Epidemiology* 26:360-365 DOI: 10.1016/j.annepidem.2016.03.004.

Vodstrcil LA, Twin J, Garland SM, Fairley CK, Hocking JS, Law MG, Plummer EL, Fethers KA, Chow EP, Tabrizi SN, and Bradshaw CS. 2017. The influence of sexual activity on
the vaginal microbiota and *Gardnerella vaginalis* clade diversity in young women. *PLoS One* 12:e0171856 DOI: 10.1371/journal.pone.0171856.

Wessels JM, Lajoie J, Vitali D, Omollo K, Kimani J, Oyugi J, Cheruiyot J, Kimani M, Mungai JN, Akolo M, Stearns JC, Surette MG, Fowke KR, and Kaushic C. 2017. Association of high-risk sexual behaviour with diversity of the vaginal microbiota and abundance of *Lactobacillus*. *PLoS One* 12:e0187612 DOI: 10.1371/journal.pone.0187612.

Woo PC, Ng KH, Lau SK, Yip KT, Fung AM, Leung KW, Tam DM, Que TL, and Yuen KY. 2003. Usefulness of the MicroSeq 500 16S ribosomal DNA-based bacterial identification system for identification of clinically significant bacterial isolates with ambiguous biochemical profiles. *Journal of Clinical Microbiology* 41:1996-2001.

Yoshimura K, Morotomi N, Fukuda K, Nakano M, Kashimura M, Hachisuga T, and Taniguchi H. 2011. Intravaginal microbial flora by the 16S rRNA gene sequencing. *American Journal of Obstetrics and Gynecology* 205:235 e231-239 DOI: 10.1016/j.ajog.2011.04.018.

Youssef N, Sheik CS, Krumholz LR, Najar FZ, Roe BA, and Elshahed MS. 2009. Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. *Applied and Environmental Microbiology* 75:5227-5236 DOI: 10.1128/AEM.00592-09.

Zhou X, Bent SJ, Schneider MG, Davis CC, Islam MR, and Forney LJ. 2004. Characterization of vaginal microbial communities in adult healthy women using cultivation-independent methods. *Microbiology* 150:2565-2573 DOI: 10.1099/mic.0.26905-0.

Zhou X, Hansmann MA, Davis CC, Suzuki H, Brown CJ, Schutte U, Pierson JD, and Forney LJ. 2010. The vaginal bacterial communities of Japanese women resemble those of women in
other racial groups. *FEMS Immunology and Medical Microbiology* 58:169-181 DOI:

10.1111/j.1574-695X.2009.00618.x.

--------------------------------------------------
Figure 1

The hierarchical clustering tree and the heat map of bacterial taxa from vaginal samples

The hierarchical clustering of 25 vaginal microbiota was generated based on the bacterial abundances of OTUs (A). The heat map showing the log2 of absolute read counts of the most abundant OTUs using the color key (B). The red- and blue bar below the tree indicates the lactobacilli-dominated and non-lactobacilli dominated group, respectively.
Manuscript to be reviewed
Figure 2

Shannon diversity index values in lactobacilli-dominated and non-lactobacilli-dominated groups

Shannon diversity index values indicating diversity of bacterial taxa in both groups, as shown using Tukey’s boxplots
Figure 3

Rarefaction curves for observed species (OTUs) (A) and Chao1 (B)

Both were used to estimate detected bacterial richness in lactobacilli-dominated and non-lactobacilli dominated groups.
Figure 4

Principle coordinates analysis (PCoA) plot of Bray-Curtis dissimilarity indices among all samples.
Figure 5

Venn diagram showing numbers of unique and shared OTUs (species) in lactobacilli-dominated and non-lactobacilli dominated groups.

- Lactobacilli-dominated: 14 (19.44%)
- Non-lactobacilli dominated: 17 (23.61%)
- 41 (56.95%)
Figure 6

The most abundant bacterial species detected in vaginal samples in the lactobacilli-dominated and non-lactobacilli dominated groups.

Data are shown as mean with SEM. The bacterial species with abundance less than 1% are not shown. These include *F. magna*, *L. kefiranofaciens*, *L. vaginalis*, *Peptoniphilus harei*, *Peptostreptococcus. anaerobius*, *S. intermedius* and *U. parvum*. * = significant difference.
Table 1 (on next page)

Sociodemographic and behavioral characteristics, and health habits of the studied subjects.
Table 1  Sociodemographic and behavioral characteristics, and health habits of the studied subjects.

|                      | Lactobacilli-dominated (N = 14) | Non-lactobacilli-dominated (N = 11) | Total (N = 25) |
|----------------------|---------------------------------|-------------------------------------|----------------|
| **n (%)**            |                                 |                                     |                |
| Median age (years) [range] | 40 [31–45]                      | 33 [27–45]                          | 39 [27–45]     |
| Median BMI (kg/m²) [range]  | 21.1 [16.0–35.8]                | 21.4 [19.3–30.4]                    | 21.2 [16.0–35.8]|
| Higher vocational or university education | 12 (85.7)                      | 10 (90.9)                           | 22 (88.0)      |
| Currently married                        | 11 (78.6)                      | 10 (90.9)                           | 21 (84.0)      |
| ≥1 child delivered                                  | 10 (71.4)                      | 6 (54.5)                            | 16 (64)        |
| Current smokers                                    | 0 (0)                          | 0 (0)                               | 0 (0)          |
| Income < 400 USD per month                          | 4 (28.6)                       | 5 (45.5)                            | 9 (36.0)       |
| Yogurt use ≥1 time per week                         | 13 (92.9)                      | 11 (100)                            | 24 (96.0)      |
| Alcohol consumption ≥1 per week                      | 3 (21.4)                       | 3 (27.3)                            | 6 (24.0)       |
| Exercise ≥1 time per week                           | 12 (85.7)                      | 8 (72.7)                            | 20 (80.0)      |
| Sexual activity ≥1 time per week                     | 9 (64.3)                       | 9 (81.8)                            | 18 (72.0)      |
| Sexual partner >1 time in previous year              | 6 (42.9)                       | 1 (9.1)                             | 7 (28.0)       |
| Regular ovulatory menstrual cycles in six months    | 12 (85.7)                      | 10 (71.4)                           | 22 (88.0)      |
| Current hormonal contraceptive use                   | 7 (50.0)                       | 2 (18.2)                            | 9 (36.0)       |
| Vaginal douching done in previous month              | 1 (7.1)                        | 1 (9.1)                             | 2 (8.0)        |
| Abnormal discharge in previous month                 | 0 (0)                          | 2 (18.2)                            | 2 (8.0)        |
| Female sterilization                                 | 5 (35.7)                       | 4 (36.4)                            | 9 (36.0)       |

Abbreviations: BMI = body mass index, USD = United States Dollar

All variables were self-reported.

Statistically not significant for any factor. P-values were calculated using chi-squared or Fisher’s exact analysis for assessment of association of frequency between groups and the Mann-Whitney U-Test for comparison of means.