Original Article
Clinical Microbiology

Ann Lab Med 2016;36:441-449
http://dx.doi.org/10.3343/alm.2016.36.5.441
ISSN 2234-3806 • eISSN 2234-3814

ANNALS OF LABORATORY MEDICINE

Analysis of the Vaginal Microbiome by Next-Generation Sequencing and Evaluation of its Performance as a Clinical Diagnostic Tool in Vaginitis

Ki Ho Hong, M.D.[1,2], Sung Kuk Hong, M.D.[2,3], Sung Im Cho, M.T.[4], Eunkyung Ra, M.T.[4], Kyung Hee Han, M.D.[5], Soon Beom Kang, M.D.[6], Eui-Chong Kim, M.D.[3,4], Sung Sup Park, M.D.[3,4], and Moon-Woo Seong, M.D.[3,4]

Department of Laboratory Medicine[1], Seoul Medical Center, Seoul, Korea; Department of Laboratory Medicine[2], College of Medicine, Seoul National University, Seoul, Korea; Department of Laboratory Medicine[3], Sheikh Khalifa Specialty Hospital, Ras Al Khaimah, United Arab Emirates; Department of Laboratory Medicine[4], Seoul National University Hospital, Seoul, Korea; Department of Obstetrics & Gynecology[5], Seoul National University Hospital Healthcare System Gangnam Center, Seoul, Korea; Department of Obstetrics & Gynecology[6], Konkuk University Medical Center, Seoul, Korea

Background: Next-generation sequencing (NGS) can detect many more microorganisms of a microbiome than traditional methods. This study aimed to analyze the vaginal microbiomes of Korean women by using NGS that included bacteria and other microorganisms. The NGS results were compared with the results of other assays, and NGS was evaluated for its feasibility for predicting vaginitis.

Methods: In total, 89 vaginal swab specimens were collected. Microscopic examinations of Gram staining and microbiological cultures were conducted on 67 specimens. NGS was performed with GS junior system on all of the vaginal specimens for the *16S rRNA*, internal transcribed spacer (*ITS*), and *Tvk* genes to detect bacteria, fungi, and *Trichomonas vaginalis*. In addition, DNA probe assays of the *Candida* spp., *Gardnerella vaginalis*, and *Trichomonas vaginalis* were performed. Various predictors of diversity that were obtained from the NGS data were analyzed to predict vaginitis.

Results: *ITS* sequences were obtained in most of the specimens (56.2%). The compositions of the intermediate and vaginitis Nugent score groups were similar to each other but differed from the composition of the normal score group. The fraction of the *Lactobacillus* spp. showed the highest area under the curve value (0.8559) in ROC curve analysis. The NGS and DNA probe assay results showed good agreement (range, 86.2-89.7%).

Conclusions: Fungi as well as bacteria should be considered for the investigation of vaginal microbiome. The intermediate and vaginitis Nugent score groups were indistinguishable in NGS. NGS is a promising diagnostic tool of the vaginal microbiome and vaginitis, although some problems need to be resolved.

Key Words: Vaginal microbiome, NGS, Vaginitis

Received: November 30, 2015
Revision received: April 3, 2016
Accepted: May 30, 2016

Corresponding author: Moon-Woo Seong
Department of Laboratory Medicine, Seoul National University Hospital, 101 Daehak-ro, Jongno-gu, Seoul 03080, Korea
Tel: +82-2-2072-4180
Fax: +82-2-747-0359
E-mail: mwseong@snu.ac.kr

INTRODUCTION

Changes in the vaginal microbiome are usually identified by clinical microbiology laboratories through microscopic examinations and vaginal swab cultures. Recently, high-throughput molecular methods, including next-generation sequencing (NGS), have been applied in investigations of the vaginal microbiome, and they have revealed that the vaginal microbiome is more complex than previously thought from the microscopic examination data.

Several studies have used NGS to analyze vaginal microbiomes in various ethnic groups, including African [1, 2], Chinese

[3, 4], and four different American ethnic groups [5]. All of these studies analyzed a number of bacterial genes, including *16S rRNA* [1, 3-5] and *cpn60* [2]. However, none of them analyzed microorganisms other than bacteria, such as fungi. Fungi were not even considered in a study of women with vulvovaginal candidiasis [4]. Although some studies have compared NGS results with those of gram-stained swabs [1, 3, 5], no studies have compared NGS results with those of microbiological cultures.

In this study, the vaginal microbiomes of Korean women were analyzed by NGS that included both bacteria and fungi. The potential for using NGS data to predict vaginitis was estimated, and the usefulness of NGS data as a clinical diagnostic tool compared with data from microbiological cultures and DNA probe assays was investigated.

METHODS

1. Specimens and culture

Eighty-nine vaginal swab specimens were collected from 87 patients who visited the gynecological clinics of Seoul National University Hospital, Seoul, Korea, from December 2011 to March 2012. Sixty-seven specimens were also subjected to Gram stain and microbiological cultures. The Nugent scores of 67 specimens were calculated on the basis of the Gram stain results [6, 7]. The specimens were grouped into three categories according to their Nugent scores: normal (score, 0-3), intermediate (score, 4-6), or vaginitis (score, ≥7). After the specimens were inoculated into blood agar plates, MacConkey agar plates, and modified Thayer-Martin agar plates, they were incubated at 37°C with 5% CO_2 for two days. The cultured organisms were identified with VITEK 2 ID Cards (bioMérieux SA, Marcy-l'Etoile, France) and MicroScan Pos ID Panels (Beckman Coulter, Inc., Brea, CA, USA). This study was approved by the Institutional Review Board of Seoul National University Hospital (review number H-1510-073-711).

2. Next-generation sequencing

The PrepMan ultra sample preparation reagent (Thermo Fisher Scientific Inc., Waltham, MA, USA) was used for the nucleic acid extraction according to the manufacturer's instructions. The sequencing of the bacteria, fungi, and *Trichomonas vaginalis* was performed on the extracted nucleic acids. For the bacteria, the V3-V5 regions of the *16S rRNA* gene were the targets. Primers 357F and 926R with an expected amplicon size of 570 bp were used for the bacterial identification [8]. For the fungi, the internal transcribed spacer (*ITS*) gene was the target. Primers ITS-5 and ITS-4 were used, and they yielded a 700-bp amplicon [9]. For *Trichomonas vaginalis*, the *Tvk* gene was the target. Primers TVK3 and TVK7 with an expected 263-bp amplicon were used [10]. The GS Junior System (454 Life Sciences, Branford, CT, USA) was used for the NGS according to the manufacturer's instructions. The minimum sequence length was 150 bp, and the minimum exponential quality score was 20. Only those sequences that fulfilled a minimum of two criteria were included in subsequent analysis.

3. Analysis pipeline

The Usearch for Windows (version 6.0.203) software was used for the sequence processing, clustering, and removing of the chimeric sequences [11]. The similarity threshold for the clustering of two sequences into the same operational taxonomic unit (OTU) was 0.97. For the removal of the chimeric sequences, both *de novo* and reference modes were used. Two open source chimeric sequence databases for the bacteria and fungi were used in the reference mode [12]. There is no known chimeric sequence database for *Trichomonas vaginalis*. When a sequence was identified as a chimeric sequence in both *de novo* and reference modes of Usearch, the sequence was regarded as a true chimeric sequence, and it was removed from the OTU.

The online Ribosomal Database Project (RDP; version 10.3.2) was used for the matching and aligning of the *16S rRNA* sequences [13]. Among the RDP sequences, near-full-length (≥ 1,200 bp) sequences of good quality were used for the matching. After comparison with the RDP sequences, each OTU was matched with the one bacterial taxon that showed the highest similarity score. If the similarity score was low (<0.5), the sequence was analyzed again with the nucleotide Basic Local Alignment Search Tool (BLAST) database [14]. The *ITS* and *Tvk* sequences were also analyzed in BLAST because these sequences are not included in the RDP [15]. PermutMatrix version 1.9.3 was used to draw a heat map for visualization of the taxonomy [16]. The Shannon diversity index was calculated to estimate the alpha diversity [17]. Two Shannon diversity indices were calculated for each specimen. First, both the *16S rRNA* and *ITS* sequences were included in the calculation. Second, only the *16S rRNA* sequences were included in the calculation.

4. DNA probe assay

The BD Affirm VPIII microbial identification test (Becton, Dickinson and Company, Franklin Lakes, NJ, USA) is a direct specimen DNA probe-based diagnostic test for the detection of the

http://dx.doi.org/10.3343/alm.2016.36.5.441

Candida spp., *Gardnerella vaginalis*, and *Trichomonas vaginalis*. VPIII was performed according to the manufacturer's instructions. In total, 87 specimens were tested with a DNA probe assay. Two specimens could not be tested because of their small volumes.

5. Comparison of the NGS, DNA probe assay, and culture data

We performed a ROC curve analysis of the various predictors of diversity in order to estimate their diagnostic value for vaginitis. Three criteria were used to compare the diagnostic criteria for vaginitis. First, a Nugent score of 4 or more was considered to indicate vaginitis, and this cutoff included both the intermediate and vaginitis groups of the original Nugent criteria [6]. Second, a Nugent score of 7 or more was considered to indicate vaginitis, and this cutoff was similar to that of the original criteria [7]. Third, microbiological culture results other than the normal vaginal flora or *Lactobacillus* spp. were considered to indicate vaginitis. Because the Nugent criteria are based on bacterial morphotypes and do not consider yeast morphotypes, we compared various other predictors, including the Shannon diversity index of *16S rRNA* and the Shannon diversity index of both *16S rRNA* and *ITS*. Other parameters, such as the total number of taxa and the fraction of *Lactobacillus* spp., were also compared.

6. Statistical analysis

Stata version 13.1 (StataCorp LP, College Station, TX, USA) was used for the statistical analysis. Paired t-tests and one-way ANOVA tests were performed to compare the Shannon diversity indices among the specimens. The Pearson's chi-square values and Cohen's kappa indices were calculated to compare the NGS data, DNA probe assays, and microbiological culture results. A ROC curve analysis was performed to evaluate the various predictors of vaginitis. *P* values less than 0.05 were considered statistically significant.

RESULTS

1. Read statistics

After removal of the chimeric sequences, 202,958 reads of the *16S rRNA* gene and 7,600 reads of the *ITS* gene were obtained from the 89 specimens (See Supplemental Data Table S1). *ITS* sequences were detected in 56.2% (50/89) of the specimens. In 38 specimens, fewer than 10 sequences were found. A *Tvk* gene read was not obtained in any specimen. We tested the NGS of the *Tvk* gene from cultured *Trichomonas vaginalis*, and the NGS detected *Trichomonas vaginalis* correctly. Therefore, we concluded that *Trichomonas vaginalis* did not exist in any of the specimens. The average sizes of the reads were 364 bp (range, 151-580) for the *16S rRNA* gene and 322 bp (range, 180-566) for the *ITS* gene.

2. OTU statistics and taxonomic allocation

1) *16S rRNA* gene

After the clustering, 202,958 reads of the *16S rRNA* sequences were clustered into 3,259 OTUs. After the matching of the *16S rRNA* sequences with those in the RDP, each OTU was allocated into the single taxon with the highest similarity. Fifty-one sequences with similarity scores less than 0.5 were rematched with the nucleotide BLAST database. Twenty-one sequences did not match any sequence with more than 97% percent identity. We excluded these sequences from further evaluation. The BLAST results for the remaining 30 sequences are shown in Supplemental Data Table S2. Of these, 20 sequences were identified as human DNA by BLAST, and these sequences were excluded from further evaluation. The 10 remaining sequences that were identified by BLAST were included in the subsequent analysis. Finally, 645 OTUs were identified to the species level, and 2,451 OTUs were identified to the genus level.

2) *ITS* gene

After the clustering, 7,600 reads of the *ITS* sequences were clustered into 112 OTUs. After taxonomic allocation, 77 OTUs were identified to the species level, and 29 OTUs were identified to the genus level. The *Candida* spp. had the highest total number of reads, whereas *Phialemonium curvatum* was detected in the highest number of specimens (See Supplemental Data Tables S3 and S4).

In both the *16S rRNA* and *ITS* taxonomies, the average read length for the taxonomic group with a low similarity score (*16S rRNA*) or low percent identity (*ITS*) was similar to that of the taxonomic group with a high similarity scores or high percent identity. Therefore, the similarity scores and percent identities appeared to be independent of read length. Similarly, the read lengths did not differ by taxonomic level.

3. Composition of the vaginal microbiomes

The compositions of the 89 specimens were clustered by Euclidean distance, as shown in the heat map in Fig. 1. The figure contains microorganisms that represent a fraction of the sequences that were greater than 0.1% of the total reads of each specimen. The 67 specimens with Nugent scores were categorized into normal, intermediate, or vaginitis groups. The patterns

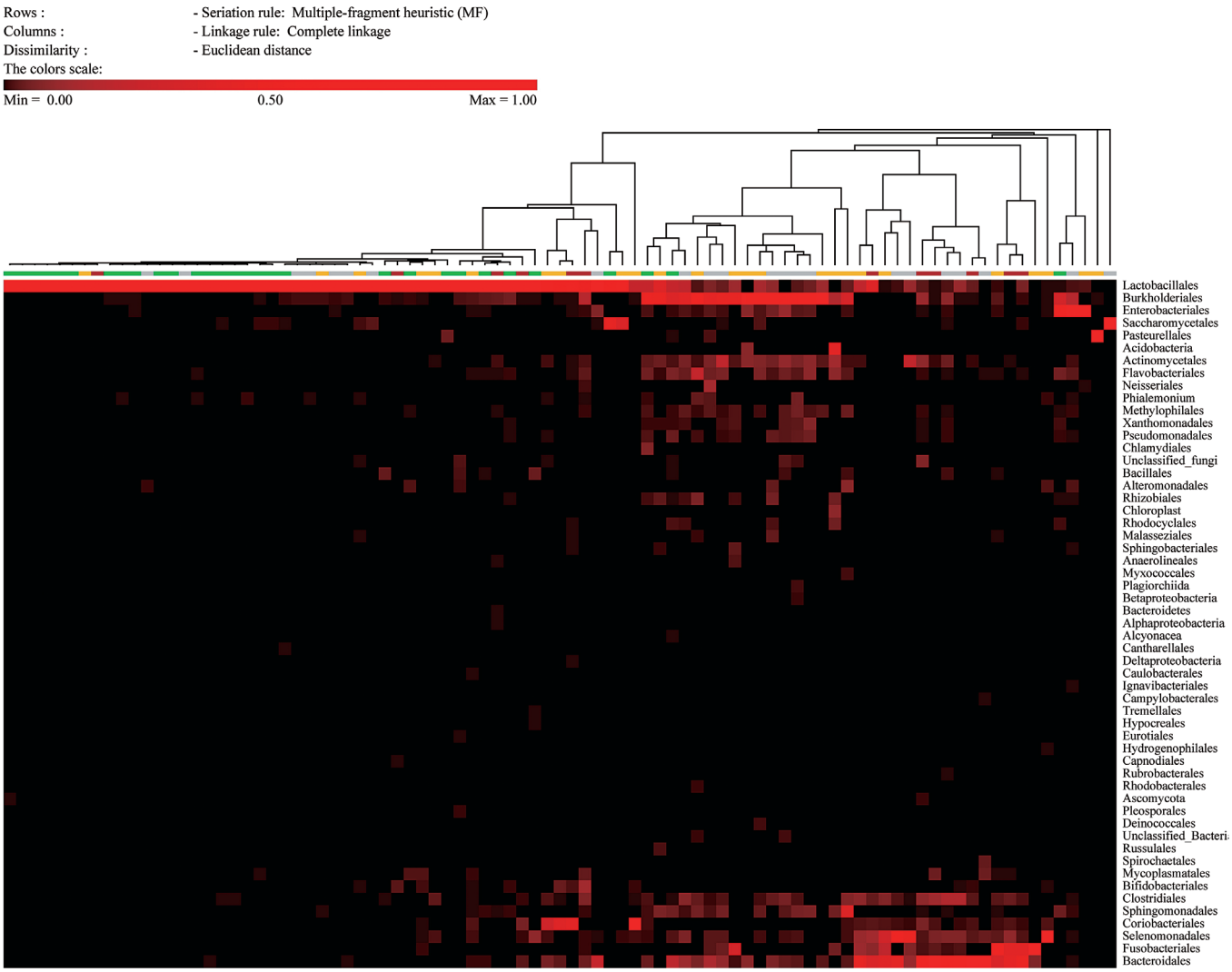

Fig. 1. Composition of the microbiomes in the 89 specimens. Each line shows the taxonomy at the order level, and each column shows a single specimen. The colored bar below the tree indicates the Nugent score group of the 67 specimens with the Gram stain results. Green, normal group; Yellow, intermediate group; Red, vaginitis group; Gray, samples without Nugent score data.

of the normal groups (green bar in Fig. 1) were distinct from the patterns of the intermediate groups (yellow bar in Fig. 1) and vaginitis groups (red bar in Fig. 1). The most abundant taxa of the three groups are shown in Table 1. In the normal group, the major taxon of the normal group was *Lactobacilliales*, and other taxa were relatively rare. The patterns of the intermediate group and the vaginitis group were similar. The genera that were more common in the intermediate and vaginitis groups than in the normal group included *Prevotella*, *Sneathia*, *Aerococcus*, *Atopobium*, *Megasphaera*, and *Cupriavidus*.

4. Diversity calculation

The mean Shannon diversity index including both the *16S rRNA* and *ITS* sequences was 1.4137 (95% confidence interval [CI]: 1.2414-1.5859) and that including the *16S rRNA* sequence only was 1.3792 (95% CI: 1.2053-1.5530). A paired t-test was used to determine a statistically significant difference (P=0.0005).

5. Association between the diversity parameters and the Nugent scores

The Shannon diversity index including the *16S rRNA* and *ITS* showed significant associations with the groups based on the Nugent scores (P=0.0037, Fig. 2A). With Bonferroni correction, there was a significant association between the normal Nugent score group and the vaginitis Nugent score group (P=0.033).

http://dx.doi.org/10.3343/alm.2016.36.5.441

Table 1. The most abundant taxa in the three Nugent score groups

Normal group (n=30)		Intermediate group (n=25)		Vaginitis group (n=12)	
Lactobacillus	83.41	*Lactobacillus*	25.19	*Lactobacillus*	38.98
Streptococcus	4.90	*Cupriavidus*	13.67	*Prevotella*	27.80
Diaphorobacter	2.50	*Sneathia*	8.65	*Sneathia*	7.48
Enterobacteriaceae	1.97	*Streptococcus*	8.27	*Aerococcus*	5.62
Candida	1.54	*Prevotella*	6.04	*Atopobium*	4.46
Cupriavidus	1.36	*Atopobium*	5.87	*Megasphaera*	1.72
Prevotella	0.80	*Megasphaera*	5.28	*Diaphorobacter*	1.67
Cloacibacterium	0.43	*Enterobacteriaceae*	4.37	*Gardnerella*	1.36
Veillonella	0.34	*Haemophilus*	3.99	*Porphyromonas*	1.29
Chlamydia	0.22	*Diaphorobacter*	2.90	*Dialister*	1.05
Comamonas	0.20	*Aerococcus*	2.17	*Cupriavidus*	1.03
Novosphingobium	0.18	*Gp4*	1.60	*Saccharofermentans*	0.94
Staphylococcus	0.16	*Sphingomonas*	1.48	*Peptoniphilus*	0.69
Haemophilus	0.14	*Candida*	1.47	*Mobiluncus*	0.69
Gemella	0.13	*Cloacibacterium*	1.23	*Anaerococcus*	0.55
Pseudomonas	0.11	*Saccharofermentans*	0.73	*Epicoccum*	0.50
Acinetobacter	0.10	*Corynebacterium*	0.55	*Coriobacteriaceae*	0.38
Alishewanella	0.09	*Novosphingobium*	0.55	*Mycoplasma*	0.35
Sphingobium	0.08	*Alishewanella*	0.38	*Moryella*	0.35
Dechloromonas	0.08	*Propionibacterium*	0.37	*Fusobacterium*	0.30

The values are the average percentages of the corresponding taxa in each Nugent group. The lowest taxonomic level is genus.

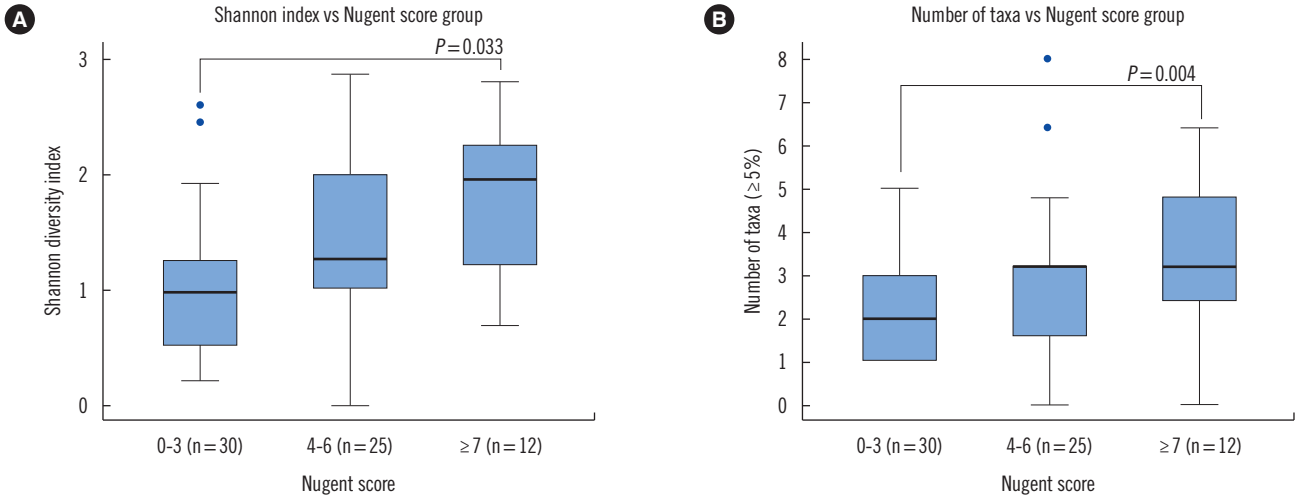

Fig. 2. Shannon diversity indices and the number of taxa (more than 5%) according to the Nugent score group. (A) The Shannon diversity indices according to the Nugent score group are shown in the form of Tukey's boxplots. The Shannon diversity index, including both bacteria and fungi, was calculated for each sample. In total, 67 Shannon diversity indices were classified into three groups according to the Nugent score of the specimen. (B) The total number of taxa (more than 5%) according to the Nugent score group are shown in the form of Tukey's boxplots.

The number of taxa representing more than 5% of the reads significantly differed among the Nugent score groups (Fig. 2B, P=0.0163). With Bonferroni correction, there was a significant association between the normal Nugent score group and the vaginitis Nugent score group (P=0.004). A one-way ANOVA showed that the proportions of *Lactobacillus* spp. differed signif-

Table 2. Comparison of the results of next-generation sequencing (NGS), DNA probe assay, and microbiological culture for the detection of *Candida* spp. and *Gardnerella vaginalis*

Method 1	Method 2	N	Organisms	SPA (%)	*P* value	κ
NGS	Culture	67	*Candida* spp.	88.1	0.003	0.363
NGS	Culture	67	*Gardnerella vaginalis*	86.6	0.003	0.509
DNA probe	Culture	67	*Candida* spp.	95.5	<0.001	0.776
DNA probe	Culture	67	*Gardnerella vaginalis*	76.1	<0.001	0.335
NGS	DNA probe	87	*Candida* spp.	86.2	<0.001	0.460
NGS	DNA probe	87	*Gardnerella vaginalis*	89.7	<0.001	0.690

The *P* values from Pearson's chi-square analysis and Cohen's kappa index were calculated for each comparison.
Abbreviations: κ, Cohen's kappa index; SPA, simple percent agreement.

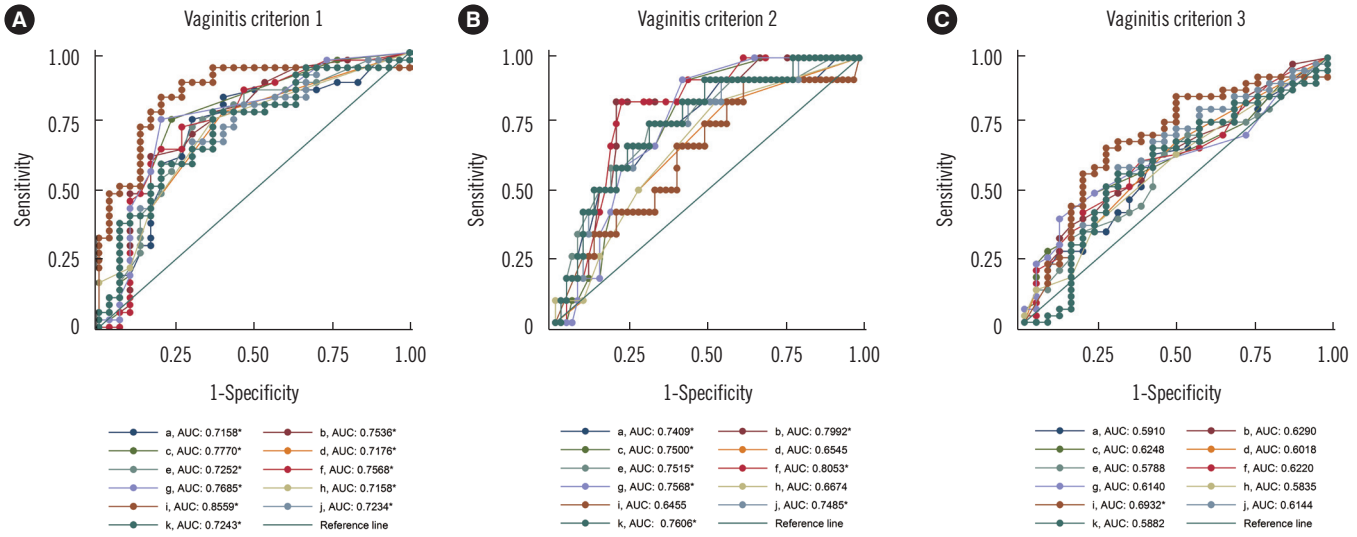

Fig. 3. ROC curves of the 11 predictors of diversity and three vaginitis criteria. The ROC curves of the 11 predictors are shown according to (A) vaginitis criterion 1, (B) vaginitis criterion 2, and (C) vaginitis criterion 3.
*$P<0.05$.
(1) Vaginitis criteria
Vaginitis criterion 1: vaginitis when the Nugent score ≥4.
Vaginitis criterion 2: vaginitis when the Nugent score ≥7.
Vaginitis criterion 3: vaginitis otherwise the culture results are either normal flora or *Lactobacillus* spp.
(2) Predictors
The specimen indicated vaginitis when the fraction of that taxon in total reads, including both the *16S rRNA* and *ITS* genes, was more than zero: (a), 0.1% (b), 1% (c), and 5% (d).
The specimen indicated vaginitis when the fraction of that taxon in total reads, including only the *16S rRNA* gene, was more than zero: (e), 0.1% (f), 1% (g), and 5% (h).
(i): The fraction of *Lactobacillus* spp. in the specimen.
(j)-(k): Shannon diversity index of the specimen, when the index was calculated from both the *16S rRNA* gene and the *ITS* gene (j) or was calculated from the *16S rRNA* gene only (k).
Abbreviations: AUC, area under the curve; *ITS*, internal transcribed spacer; rRNA, ribosomal RNA.

icantly among the Nugent score groups ($P<0.0001$), and the proportions did not increase in accordance with the Nugent score group. The mean proportions of *Lactobacillus* spp. in the normal, intermediate, and vaginitis groups were 0.827, 0.262, and 0.426, respectively.

6. Comparison of the NGS, DNA probe assay, and culture data

The simple percent agreements of the NGS, DNA probe assay, and culture data ranged from 76.1% to 95.5% for the detection of *Candida* spp. and *Gardnerella vaginalis* (Table 2). Pearson's chi-square analysis showed significant associations for all three assays. It was difficult to directly compare the NGS and culture results because the NGS detected many microorganisms and

http://dx.doi.org/10.3343/alm.2016.36.5.441

the cultures usually detected only a few. In our study, the NGS and culture results were considered in complete agreement if the NGS results included all of the cultured microorganisms. With this definition of complete agreement, the NGS and culture data showed 73.1% agreement (49/67). Two specimens were considered in partial agreement. In one specimen showing partial agreement, *Enterococcus faecalis* and *Candida albicans* were grown, whereas the NGS detected only *Candida albicans*. In the other specimen, *Escherichia coli* and *Gardnerella vaginalis* were grown in culture, whereas the NGS detected only *Gardnerella vaginalis*. The remaining specimens with discordant results are shown in Supplemental Data Table S5.

7. Comparison of the various predictors of diversity for the diagnosis of vaginitis

The highest area under the curve (AUC), which was 0.8559, was obtained on the basis of the fraction of lactobacilli and the first criterion for vaginitis (Fig. 3A). When this parameter and criterion combination was applied with a 12.45% lactobacilli fraction cutoff, the sensitivity was 83.78% (95% CI: 68.0-93.8%) and the specificity was 80.00% (95% CI: 61.4-92.3%). All of the predictors showed significant AUC values with the first vaginitis criterion (Nugent score: ≥4).

DISCUSSION

Several studies have analyzed the vaginal microbiome with NGS technology [1-5]. Hummelen *et al* [1] studied the vaginal microbiota of 132 HIV-positive Tanzanian women. Martin *et al* [18] studied the vaginal swabs of 92 American women (ethnicity unknown). This study showed the highest average read length of *16s rRNA* sequences of all of the earlier studies (480 bp). However, the previous studies analyzed only bacteria in the vaginal microbiome and did not analyze fungi. As many bacterial vaginitis cases are accompanied by vulvovaginal candidiasis, an analysis of fungi may provide more information about the vaginal microbiome. In our study, the *ITS* sequence was found in 56.2% of the specimens, although many specimens included only a few reads of this gene. The Shannon diversity indices also showed significant differences when fungi were included. These results suggest that fungi, as well as bacteria, should be considered in investigations of the vaginal microbiome.

Ravel *et al* [5] reported a significant difference in the vaginal microbiome composition among four ethnic groups (white, black, Hispanic, and Asian). In that study, the most common group of bacteria that was detected in the asymptomatic women with Asian ethnicity was *Lactobacillus iners* (42.7%), and the second most common group was *L. crispatus* (25%). We analyzed the vaginal microbiome of Korean women with NGS, but our data could not be compared directly with the results of Ravel's study because many of the lactobacilli in our study were identified at only the genus level (63.0% in the normal Nugent score group). However, *L. iners* was more common than *L. crispatus* among the *Lactobacillus* identified at the species level in the normal Nugent score group (14.7% vs 11.7%).

The microbiota of the normal Nugent score group were mainly comprised of *Lactobacillus*, as was expected on the basis of previous reports (Fig. 1). The Shannon diversity index, the number of species, and the proportion of *Lactobacillus* spp. differed significantly among the Nugent score groups (Figs. 2 and 3). However, the compositions of the intermediate Nugent score group and vaginitis Nugent score group were similar. One cause of this finding could be that the intermediate and vaginitis groups might have been on the same clinical spectrum and thus may have shared similar microbiome patterns. In Fig. 1, the *Lactobacillus* fraction showed associations with the Nugent score when only the difference between the normal group and the other two groups (combination of the intermediate and vaginitis groups) was considered. However, there was no proportional increase in the fraction according to the Nugent score group (Table 1). Other possible causes included bias by the examiner who graded the Nugent scores and the relatively small number of specimens in the vaginitis group. In our data, all 12 of the specimens comprising the vaginitis group (Nugent score ≥7) had a Nugent score of 8, which suggested the possibility of bias when the Nugent scores were assigned. Because the Nugent scores were greatly dependent on the examiner, bias was possible.

Among the various predictors of diversity, the *Lactobacillus* fraction showed the highest AUC in the ROC curve analysis (Fig. 3). Because the vaginitis criteria that were used relied on bacterial morphotypes, the predictors that were based on *16S rRNA* would be expected to be in good agreement with the vaginitis criteria. If other criteria that consider bacteria and fungi are developed, other predictors may show higher correlations. Because predictors from the NGS data are much less affected by potential examiner subjectivity than the Nugent score is, they are more objective for predicting vaginal microbiome diversity and the associated clinical conditions.

We compared the NGS, DNA probe assay, and microbiological culture data for the detection of vaginal microorganisms. Few studies have compared the use of NGS and culture in in-

vestigations of vaginal microbiomes.

Smidt *et al* [19] compared NGS, quantitative PCR, and culture-based methods in the identification of *Lactobacillus* spp. To the best of our knowledge, this is the only prior study that compared culture and NGS in an investigation of vaginal specimens. They compared the species agreement within the *Lactobacillus* genus and reported general concordance of the three methods in the detection of *L. crispatus*, *L. jensenii*, and *L. gasseri* but not of *L. iners*. In our study, the NGS and microbiological culture showed only 73.1% agreement, whereas the NGS and DNA probe assay showed good agreement.

Salipante *et al* [20] reported similar results on sputum specimens from cystic fibrosis patients. In their study, 17.3% of the pathogens were identified only by culture, and the total agreement between the NGS and culture results was 56.7%. Some of the reads were lost during the denoising steps of NGS and that this caused some of the failures. They also suggested that the discrepancies reflect various factors, including inefficient DNA extraction from particular organisms, primer bias, and/or properties of the specimens themselves, including internal sample heterogeneity.

In addition, Toma *et al* [21] reported discrepancies between the culture and NGS results for endotracheal aspirates. Interestingly, they used multiple databases, and the discrepancy percentages differed according to database. The NGS and culture results coincided in 85% of the samples using three databases. They suggested that short microbial reads and amplification bias resulting from mismatches of the universal primers in some specific bacteria might have caused such discrepancies. We did not identify any studies that reported discrepancies between the NGS and culture results in investigations of the vaginal microbiome, but the factors mentioned above could have had similar effects in our study.

The RDP is a very convenient tool for analysis of the *16S rRNA* sequences in the NGS data. It provides a taxonomic ranking of sequences in a form that can easily be converted to various data platforms. However, our results showed some possible significant mismatches with low similarity scores. Fettweis *et al* [22] noted similar mismatches. In our study, several RDP results with low similarity scores were identified as human DNA and *Candida* spp. Because vaginal swabs can include nucleic acid from humans and microorganisms other than bacteria, sequences with poor similarity scores should be analyzed with other databases.

This study had some limitations. First, we could not evaluate the various analytical performance parameters, but we investigated the possibility of the use of NGS as a clinical diagnostic tool. We evaluated the accuracy of the NGS assay by comparison with the culture and DNA probe assay results. Second, the NGS results and microbiological culture results were indirectly compared.

Other problems remain with the use of NGS as a clinical diagnostic tool [20]. Although the cost of NGS is decreasing rapidly, it is still too high for use as a clinical test. The interpretation of the NGS data is complicated, and few recognizable standards for NGS interpretation exist.

In our study, NGS data elucidated the frequent existence of fungi in vaginal specimens and the similarity of composition between the intermediate and vaginitis Nugent score groups. Although some problems remain to be resolved, NGS could become a powerful method for investigating vaginal microbiomes in clinical diseases that are associated with specific vaginal microbiome profiles.

Authors' Disclosures of Potential Conflicts of Interest

No potential conflicts of interest relevant to this article were reported.

REFERENCES

1. Hummelen R, Fernandes AD, Macklaim JM, Dickson RJ, Changalucha J, Gloor GB, et al. Deep sequencing of the vaginal microbiota of women with HIV. PLoS One 2010;5:e12078.
2. Schellenberg JJ, Links MG, Hill JE, Dumonceaux TJ, Kimani J, Jaoko W, et al. Molecular definition of vaginal microbiota in East African commercial sex workers. Appl Environ Microbiol 2011;77:4066-74.
3. Ling Z, Kong J, Liu F, Zhu H, Chen X, Wang Y, et al. Molecular analysis of the diversity of vaginal microbiota associated with bacterial vaginosis. BMC Genomics 2010;11:488.
4. Liu MB, Xu SR, He Y, Deng GH, Sheng HF, Huang XM, et al. Diverse vaginal microbiomes in reproductive-age women with vulvovaginal candidiasis. PLoS One 2013;8:e79812.
5. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, et al. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci U S A 2011;108 (S1):S4680-7.
6. Nugent RP, Krohn MA, Hillier SL. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. J Clin Microbiol 1991;29:297-301.
7. Rodrigues FS, Peixoto S, Adami F, Alves Bda C, Gehrke Fde S, Azzalis LA, et al. Proposal of a new cutoff for Nugent criteria in the diagnosis of bacterial vaginosis. J Microbiol Methods 2015;115:144-6.
8. Jumpstart Consortium Human Microbiome Project Data Generation Working Group. 16S 454 Sequencing Protocol. http://www.hmpdacc.org/doc/16S_Sequencing_SOP_4.2.2.pdf (Updated on Oct 27, 2010).
9. CLSI. Interpretive criteria for identification of bacteria and fungi by DNA target sequencing; Approved guideline. CLSI document MM18-A.

http://dx.doi.org/10.3343/alm.2016.36.5.441

Wayne, PA: Clinical and Laboratory Standards Institute, 2008.
10. Lawing LF, Hedges SR, Schwebke JR. Detection of trichomonosis in vaginal and urine specimens from women by culture and PCR. J Clin Microbiol 2000;38:3585-8.
11. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 2010;26:2460-1.
12. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 2011; 27:2194-200.
13. Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, et al. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. Genome Res 2011;21:494-504.
14. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 2007;73:5261-7.
15. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403-10.
16. Caraux G and Pinloche S. PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order. Bioinformatics 2005;21:1280-1.
17. Shannon CE. A mathematical theory of communication. Bell System Tech J 1948;27:379-423.
18. Martin DH, Zozaya M, Lillis R, Miller J, Ferris MJ. The microbiota of the human genitourinary tract: trying to see the forest through the trees. Trans Am Clin Climatol Assoc 2012;123:242-56.
19. Smidt I, Kiiker R, Oopkaup H, Lapp E, Rööp T, Truusalu K, et al. Comparison of detection methods for vaginal lactobacilli. Benef Microbes 2015;6:747-51.
20. Salipante SJ, Sengupta DJ, Rosenthal C, Costa G, Spangler J, Sims EH, et al. Rapid 16S rRNA next-generation sequencing of polymicrobial clinical samples for diagnosis of complex bacterial infections. PLoS One 2013;8:e65226.
21. Toma I, Siegel MO, Keiser J, Yakovleva A, Kim A, Davenport L, et al. Single-molecule long-read 16S sequencing to characterize the lung microbiome from mechanically ventilated patients with suspected pneumonia. J Clin Microbiol 2014;52:3913-21.
22. Fettweis JM, Serrano MG, Sheth NU, Mayer CM, Glascock AL, Brooks JP, et al. Species-level classification of the vaginal microbiome. BMC Genomics 2012;13(S 8):S17.