AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Clinical Microbiology

Development and Validation of a Highly Accurate Quantitative Real-Time PCR Assay for Diagnosis of Bacterial Vaginosis

David W. Hilbert,[a] William L. Smith,[a] Sean G. Chadwick,[a] Geoffrey Toner,[a] Eli Mordechai,[a] Martin E. Adelson,[a] Tina J. Aguin,[b] Jack D. Sobel,[c] Scott E. Gygax[a]

Femeris Women's Health Research Center, Medical Diagnostic Laboratories, a member of Genesis Biotechnology group, Hamilton, New Jersey, USA[a]; Department of Obstetrics and Gynecology, Wayne State University School of Medicine, Detroit, Michigan, USA[b]; Division of Infectious Diseases, Wayne State University School of Medicine, Detroit, Michigan, USA[c]

Bacterial vaginosis (BV) is the most common gynecological infection in the United States. Diagnosis based on Amsel's criteria can be challenging and can be aided by laboratory-based testing. A standard method for diagnosis in research studies is enumeration of bacterial morphotypes of a Gram-stained vaginal smear (i.e., Nugent scoring). However, this technique is subjective, requires specialized training, and is not widely available. Therefore, a highly accurate molecular assay for the diagnosis of BV would be of great utility. We analyzed 385 vaginal specimens collected prospectively from subjects who were evaluated for BV by clinical signs and Nugent scoring. We analyzed quantitative real-time PCR (qPCR) assays on DNA extracted from these specimens to quantify nine organisms associated with vaginal health or disease: *Gardnerella vaginalis*, *Atopobium vaginae*, BV-associated bacteria 2 (BVAB2, an uncultured member of the order *Clostridiales*), *Megasphaera* phylotype 1 or 2, *Lactobacillus iners*, *Lactobacillus crispatus*, *Lactobacillus gasseri*, and *Lactobacillus jensenii*. We generated a logistic regression model that identified *G. vaginalis*, *A. vaginae*, and *Megasphaera* phylotypes 1 and 2 as the organisms for which quantification provided the most accurate diagnosis of symptomatic BV, as defined by Amsel's criteria and Nugent scoring, with 92% sensitivity, 95% specificity, 94% positive predictive value, and 94% negative predictive value. The inclusion of *Lactobacillus* spp. did not contribute sufficiently to the quantitative model for symptomatic BV detection. This molecular assay is a highly accurate laboratory tool to assist in the diagnosis of symptomatic BV.

Bacterial vaginosis (BV) is the most common gynecological infection in the United States, affecting 29% of women (1). The disease is characterized by a shift in the vaginal flora from commensal lactobacillus to diverse fastidious Gram-negative and variable anaerobic and facultative species (2). Symptomatic disease is characterized by vaginal discharge that may or may not have a fishy odor; however, >50% of cases are asymptomatic (1, 3). The increased risk of individuals with BV for pregnancy complications or sexually transmitted infection (STI) acquisition is present in both symptomatic and asymptomatic individuals (4–6). BV is associated with preterm labor in pregnant women (7), postoperative infections after hysterectomy (8) or surgical abortion (9), and acquisition of STIs, including HIV and *Trichomonas vaginalis* (5, 10). Diagnosis of symptomatic BV is based upon the presence of ≥3 of the following clinical signs (Amsel's criteria): (i) elevated vaginal pH (>4.5), (ii) a homogenous thin gray-white vaginal discharge, (iii) the presence of clue cells (bacteria-covered exfoliated vaginal epithelial cells) by wet-mount microscopy of a vaginal smear, or (iv) a positive whiff test (fishy odor after addition of 10% KOH to a sample of vaginal discharge) (3). Diagnosis based upon clinical signs without laboratory testing has 90% sensitivity but only 77% specificity (11). Therefore, characterization of the vaginal flora by enumeration of bacterial morphotypes after Gram staining of a vaginal smear (i.e., Nugent scoring) is considered the standard method for BV diagnosis in research studies (12). Although highly accurate, this technique requires specialized training and is subjective, and microscopic evaluation is not widely available for clinicians to assist in the diagnosis of BV. Thus, development of a highly accurate molecular assay to aid in the diagnosis of BV would be of significant clinical utility.

BV has traditionally been associated with *Gardnerella vaginalis* since the initial description of *G. vaginalis* in 1955 as associated with the disease (13). Our understanding of the etiology of BV became more complex as it was later associated with culturing of diverse fastidious anaerobic species, such as *Bacteroides*, *Mobiluncus*, *Prevotella*, and *Atopobium* (14, 15). DNA sequencing of vaginal specimens subsequently identified a variety of uncultured bacteria associated with BV, such as BV-associated bacteria (BVAB)-1, BVAB-2, or BVAB-3 (which are uncultured members of the order *Clostridiales*), *Megasphaera*, *Eggerthella*, and *Leptotrichia* (16). A number of prior studies have evaluated PCR assays for BV diagnosis (15, 17, 18) and obtained promising results. We have expanded upon these studies by taking advantage of a newly developed multiplex real-time quantitative PCR (qPCR) assay for the quantification of four *Lactobacillus* spp. (19): *Lactobacillus crispatus*, *Lactobacillus jensenii*, and *Lactobacillus gasseri*, which are commensal species associated with vaginal health and are depleted in BV (20), and *Lactobacillus iners*, which is found in healthy individuals and in those with BV (21). We have used this assay in

Received 4 December 2015 **Returned for modification** 29 December 2015
Accepted 22 January 2016

Accepted manuscript posted online 27 January 2016

Citation Hilbert DW, Smith WL, Chadwick SG, Toner G, Mordechai E, Adelson ME, Aguin TJ, Sobel JD, Gygax SE. 2016. Development and validation of a highly accurate quantitative real-time PCR assay for diagnosis of bacterial vaginosis. J Clin Microbiol 54:1017–1024. doi:10.1128/JCM.03104-15.

Editor: A. B. Onderdonk

Address correspondence to David W. Hilbert, dhilbert@femeris.com, or Scott E. Gygax, sgygax@femeris.com.

TABLE 1 Characteristics of specimens evaluated in the study

Characteristic	Initial visit	Post-treatment (1 week)	Short-term follow-up (40–45 days)	Long-term follow-up (~monthly)	Total
No. of women	146	103	85	24	147[a]
No. of samples	146	103	85	51	385
No. (%) with Amsel's criteria 0–2	23 (16)	97 (94)	53 (62)	41 (80)	214 (56)
No. (%) with Nugent score 0–3[b]	16 (70)	84 (85)	41 (77)	33 (80)	174 (81)
No. (%) with Nugent score 4–6[b]	7 (30)	13 (13)	12 (23)	8 (20)	40 (19)
No. (%) with Amsel's criteria 3–4	123 (84)	6 (5.8)	32 (38)	10 (20)	171 (44)
No. (%) with Nugent score 4–6[c]	16 (13)	0 (0.0)	4 (13)	0 (0.0)	20 (12)
No. (%) with Nugent score 7–10[c]	107 (87)	6 (86)	28 (88)	10 (100)	151 (88)

[a] From one woman, specimens were evaluated at post-treatment, short-term follow-up, and long-term follow-up visits, but her initial visit specimen was excluded due to discordant Amsel's/Nugent results.
[b] Percentage represents a percent of the total specimens with 0–2 Amsel's criteria.
[c] Percentage represents a percent of the total specimens with 3–4 Amsel's criteria.

conjunction with individual qPCR assays to quantify *G. vaginalis*, *A. vaginae*, BVAB2, and *Megasphaera* phylotypes 1 and 2 on vaginal specimens collected from patients who were characterized by both Amsel's criteria and Nugent scoring for BV diagnosis. Furthermore, we performed logistic regression and receiver operator characteristic (ROC) analysis on our qPCR data to generate a highly accurate diagnostic tool to assist clinicians in the diagnosis of BV.

MATERIALS AND METHODS

Clinical specimens. A longitudinal clinical study was performed at the vaginitis clinic of Wayne State University Medical School in Detroit, MI, USA. Premenopausal women age 18 years or older were enrolled. The study took place from March 2012 to November 2013. Exclusion criteria included menstruation, pregnancy, immunocompromised status, antibiotic treatment within the past 30 days, and current treatment with intravaginal anti-inflammatory or antihistamine agents or for any non-BV urogenital infection. A total of 149 women were enrolled in the study. The average age was 34 years; 3 women did not report their ages. Race was self reported as follows: 102 women (68%) were black, 42 (28%) were Caucasian, 3 (2.0%) were Asian, 1 (0.67%) was Hispanic, and 1 (0.67%) did not report her race. Informed consent was obtained from all subjects, and human experimentation guidelines of the U.S. Department of Health and Human Services were strictly followed. Subjects were evaluated for BV by Amsel's criteria (3), and 3 to 4 clinical signs needed to be present for BV diagnosis. In addition, vaginal smears were evaluated by Nugent scoring (12), and specimens were categorized as normal (score, 0 to 3), intermediate (score, 4 to 6), or abnormal (score, 7 to 10). A total of 400 samples were collected in the study from 149 women. We excluded 15 specimens for the following reasons: absence of Nugent score ($n = 6$), no amplifiable microorganism DNA present ($n = 1$), or discordance between Nugent score and Amsel's criteria ($n = 8$). In the cases of discordance, either abnormal vaginal flora was present with 0 to 2 Amsel criteria, a condition sometimes referred to asymptomatic BV, a well-established phenomenon (1, 3, 22) ($n = 7$), or normal flora was present with 3 to 4 Amsel criteria, most likely a case of vaginitis of non-BV etiology ($n = 1$). Vaginal specimens were isolated using a OneSwab (Medical Diagnostic Laboratories) and stored in universal transport medium at room temperature (UTM-RT) (Copan Italia), which consists of Hanks' balanced salts, bovine serum albumin, L-cysteine, gelatin, sucrose, L-glutamic acid, HEPES buffer, vancomycin, amphotericin B, colistin, and phenol red (pH 7.3 ± 0.2). This transport medium has been previously validated with respect to detection of vaginal microorganisms (23). Vaginal specimens were shipped to Medical Diagnostic Laboratories, where they were stored at −80°C until analyzed.

Molecular analysis. DNA was extracted from clinical specimens using mechanical disruption and the QIAamp minikit (Qiagen), as described previously (24). qPCR assays used to quantify BV-associated microbes and *Lactobacillus* spp. have been previously described (19, 25). The *Megasphaera* spp. reaction is run as a duplex PCR using both phylotype 1 and phylotype 2 probes; the *Lactobacillus* spp. PCR is run as a multiplex, using probes for *L. crispatus*, *L. gasseri*, *L. iners*, and *L. jensenii*; the other reactions are run individually. All reactions were performed using a CFX384 real-time thermocycler (Bio-Rad). Reactions conditions were 2 min at 50°C and 2 min at 95°C, followed by 40 cycles of 10 s at 95°C and 45 s at 60°C. Final primer concentrations were 200 to 800 nM, and final probe concentrations were 75 to 200 nM. The quantity of PCR target copies detected was expressed per reaction, each one using 2.5 μl of DNA as the template in a 25-μl reaction. The concentration of targets was determined by comparing the reaction to a standard curve that was generated using 10^2 to 10^6 copies of a plasmid encoding the amplification target. A cutoff of 10 copies per reaction was used to determine if a sample was positive for a given organism. Each reaction plate was run with multiple negative controls to ensure no template DNA contamination of reagents.

Statistical analysis. Sensitivity, specificity, and likelihood ratios for diagnostic performance using categorical data were performed using MedCalc (http://www.medcalc.org/calc/diagnostic_test.php). Logistic regression and receiver operator characteristic (ROC) analysis was performed using R version 3.0.3. The MuMIn and bestglm packages were used to identify the best subset of assays for generalized linear model analysis. The predict function from the stats package was used to calculate a probability for each specimen to be indicative of BV, and these probabilities were further analyzed by ROC analysis to determine the optimal cutpoints to achieve maximum efficiency. The optimal model identified by logistic regression was as follows (all organism values are copies/reaction): −2.86 (intercept) + (*G. vaginalis* × 1.43E−7) + (*A. vaginae* × 1.87E−6) + (*Megasphaera* phylotype 1 × 6.73E−6) + (*Megasphaera* phylotype 2 × 4.50E−3). This sum is the natural log of the odds (O) of being BV positive, which can be used to determine the probability (p) of BV using the equation $p = O/(1 + O)$. These probabilities were further characterized by performing ROC analysis using the packages ROCR and OptimalCutpoints. Cutpoints were determined by maximum efficiency, which is the value at which the fewest incidents of misclassification occurred. The area under the curve (AUC) of ROC curves were compared using the DeLong test. Multiple comparisons were controlled for using the Benjamini-Hochberg procedure.

RESULTS

Characteristics of clinical specimens. We evaluated a total of 385 vaginal specimens collected from 146 women (Table 1). Our rationale for using specimens obtained from a longitudinal study is

TABLE 2 Evaluation of qPCR assays for BV diagnosis based upon nonquantitative detection of organisms

Organism	Sensitivity (%) (95% CI[a])	Specificity (%) (95% CI)	Likelihood ratio (95% CI)
G. vaginalis	98.3 (95.0–99.6)	42.5 (35.8–49.5)	1.71 (1.52–1.92)
A. vaginae	93.6 (88.8–96.7)	57.5 (50.6–64.2)	2.20 (1.87–2.58)
BVAB2	80.1 (73.3–85.8)	82.2 (76.5–87.1)	4.32 (3.23–5.76)
Megasphaera phylotype 1	68.4 (60.9–75.3)	98.1 (95.3–99.5)	36.6 (13.8–97.2)
Megasphaera phylotype 2	31.0 (24.2–38.5)	100 (98.3–100)	Infinity
Megasphaera phylotype 1 or phylotype 2	80.7 (74.0–86.3)	98.1 (95.3–99.4)	43.2 (16.3–114)
L. iners	94.2 (89.5–97.2)	17.8 (12.9–23.6)	1.14 (1.06–1.23)
Absence of *L. crispatus*	82.5 (75.9–87.8)	40.7 (34.0–47.6)	2.32 (1.61–3.33)
Absence of *L. gasseri*	94.2 (89.5–97.2)	14.5 (10.1–19.9)	1.10 (1.03–1.18)
Absence of *L. jensenii*	90.6 (85.3–94.6)	39.3 (32.7–46.1)	1.49 (1.33–1.68)
Absence of *L. crispatus*, *L. gasseri*, and *L. jensenii*	70.1 (63.5–76.1)	73.7 (66.4–80.1)	2.46 (1.97–3.08)

[a] CI, confidence interval.

that clinicians must evaluate women for BV in several different clinical settings (initial diagnosis, test-of-cure, short- and long-term follow-ups) and thus any diagnostic tool we developed had to be sufficiently robust to function in all of them. We enrolled 123 women with BV as defined by Amsel's criteria along with 23 healthy women; women with BV were treated with standard antimicrobial therapy, and 103 returned 1 week later for a posttreatment follow-up visit, of whom 97 were successfully treated and 6 had persistent BV. Of these women, 84 were evaluated at a 40- to 45-day short-term follow-up visit, at which 32 of them had recurrent BV. Finally, a total of 51 specimens were collected from 24 women at monthly intervals for up to 5 additional months (long-term follow-up). Of the 214 specimens from women without BV, 174 had normal flora (Nugent scores, 0 to 3) and 40 had intermediate flora (Nugent scores, 4 to 6). Of the 171 specimens collected from women with BV, 151 had abnormal flora (Nugent scores, 7 to 10), and 20 had intermediate flora. The presence of intermediate Nugent scores (4–6) among both BV and non-BV cohorts is consistent with prior studies that observed a range of clinical BV signs in individuals with intermediate flora (26, 27). Our goal was to develop a diagnostic tool that would identify women with BV who require treatment. Therefore, we defined BV by Amsel's criteria for this study, as current guidelines state that women with abnormal flora but no symptoms of BV do not require treatment (28).

Evaluation of qPCR assays for BV diagnosis. We generated a panel of qPCR assays to detect organisms associated with BV and to detect *Lactobacillus* spp. associated with vaginal health. *G. vaginalis* has historically been the organism most strongly associated with the disease (13, 29). *A. vaginae* is a fastidious organism, discovered later through culture of vaginal specimens, and has also been associated with BV (15, 30). DNA sequencing has identified a number of organisms, not yet cultured, that are associated with BV as well (16). We chose to focus on BVAB2 and *Megasphaera* phylotypes, as these organisms have very strong associations with BV (the odds ratios for BV among positive individuals is ≥116) (17). In addition, the onset of BV is associated with the loss of commensal *Lactobacillus* spp., such as *L. crispatus*, *L. gasseri*, and *L. jensenii* (21, 31, 32), whereas *L. iners* is an atypical *Lactobacillus* species in that it is present in healthy individuals as well as those with BV (21). Although the role of *L. iners* in BV is unclear, it is the predominant vaginal species in certain individuals (32). Therefore, we also evaluated a recently developed multiplex PCR assay, that has been used to quantify *Lactobacillus* spp. on vaginal specimens (19), in combination with the assays quantifying BV-associated organisms.

We performed these qPCR assays on DNA extracted from BV ($n = 171$) or non-BV ($n = 214$) vaginal specimens. Positivity rates (%) among BV and non-BV specimens, respectively, were as follows by BV-associated organism: *G. vaginalis* (98% and 57%), *A. vaginae* (94% and 43%), BVAB2 (80% and 18%), *Megasphaera* phylotype 1 (68% and 2%), and *Megasphaera* phylotype 2 (31% and 0%). The prevalence (%) of *L. crispatus*, *L. gasseri*, and *L. jensenii* strains among non-BV specimens (41%, 14%, and 39%, respectively) was higher than that of BV specimens (18%, 6%, and 9%, respectively). *L. iners*, very common in both cohorts, was detected in 94% of BV and 82% of non-BV specimens. We initially analyzed data categorically with respect to BV status (Table 2). *G. vaginalis*, *A. vaginae*, and BVAB2 provided low diagnostic specificity due to their high prevalences among non-BV specimens. In contrast, *Megasphaera* phylotypes 1 and 2 were highly specific yet, even in combination, provided only 81% sensitivity for BV detection. Detection of *Lactobacillus* spp. alone as negative indicators of BV has limited utility, as they provided, even in combination (excluding *L. insers*), only 70% sensitivity and 74% specificity. Therefore, detection of organisms alone, not accounting for the quantity present, was insufficient in this study to generate a highly accurate BV diagnostic test.

G. vaginalis, *A. vaginae*, and BVAB2 organisms were highly prevalent in both BV and non-BV specimens, with a far higher median concentration for the former. Therefore, we wanted to determine if we could establish a threshold for organism concentration that would allow for BV diagnosis. Indeed, a prior study has found that determining the concentration of *G. vaginalis* or *A.*

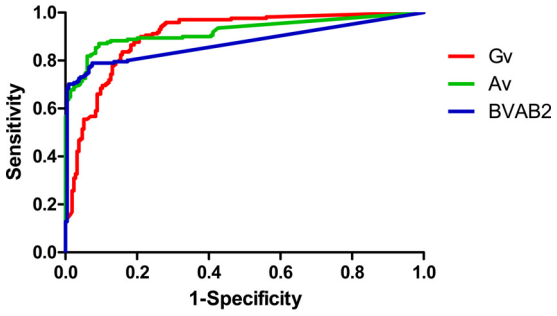

FIG 1 ROC analysis of qPCR data for individual BV-associated organisms. Av, *A. vaginae*; Gv, *G. vaginalis*.

TABLE 3 Diagnostic performance of selected qPCR assays using cutoffs determined by ROC analysis

Organism or model	AUC (95% CI)	Cutoff[a]	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Likelihood ratio (95% CI)
G. vaginalis	0.90 (0.87–0.94)	1.10E5	83.6 (77.2–88.8)	84.1 (78.5–88.7)	5.26 (3.84–7.21)
A. vaginae	0.92 (0.89–0.95)	1.10E5	87.1 (81.2–91.8)	90.7 (85.9–94.2)	9.32 (6.12–14.2)
BV	0.87 (0.84–0.91)	1.00E3	78.9 (72.1–84.8)	92.5 (88.1–95.7)	10.6 (6.55–17.0)
G. vaginalis + *A. vaginae*	0.96 (0.94–0.98)	1.26E6	87.1 (81.2–91.6)	93.9 (89.8–96.7)	14.3 (8.44–24.4)
G. vaginalis + BVAB2	0.93 (0.91–0.96)	6.41E5	85.4 (79.2–90.3)	90.7 (85.9–94.2)	9.14 (5.99–13.9)
A. vaginae + BVAB2	0.95 (0.93–0.98)	1.53E5	92.4 (87.4–95.9)	90.2 (85.4–93.8)	9.42 (6.26–14.2)
G. vaginalis + *A. vaginae* BVAB2	0.96 (0.94–0.98)	1.17E6	91.2 (85.9–95.0)	93.0 (88.7–96.0)	13.0 (7.97–21.2)
BIC[b] model	0.98 (0.97–0.99)	0.20	91.8 (86.6–95.5)	95.3 (91.6–97.7)	19.7 (10.7–36.1)

[a] Cutoff is copies/reaction except for BIC, for which the cutoff is the predicted probability of the logistic regression model, where 0 = healthy and 1 = BV.
[b] BIC, Bayesian information criterion.

vaginae in specimens can be a highly accurate BV diagnostic method (18). To this end, we performed ROC analysis on our qPCR data (Fig. 1) and determined cutoffs based on maximum accuracy (Table 3). We found that analyzing the data in this fashion dramatically improved specificity compared to our categorical analysis of these organisms (Table 2). For example, instituting a cutoff of 1.10E5 copies/reaction of the *G. vaginalis* target improved specificity from 42% to 84% while modestly reducing sensitivity from 98% to 84%. Using this same cutoff for *A. vaginae* similarly improved specificity from 58% to 91%, with a concomitant drop in sensitivity from 94% to 87%. The optimal cutoff for BVAB2 was far lower (1.00E3 copies/reaction), but we obtained a modest increase in specificity (82% to 92%) with essentially the same sensitivity (79% to 80%). After controlling for multiple comparisons, the only difference between these organisms was that the AUC for the ROC of *A. vaginae* was greater than that of BVAB2 ($P = 0.025$) (Table 4).

We next evaluated the combined titers of *G. vaginalis*, *A. vaginae*, and BVAB2. ROC analysis (Fig. 2) revealed that these combinations generated a greater AUC than when individual organisms were used (Table 3). The combination of *G. vaginalis* and *A. vaginae* improved specificity to 94% from 84% to 87% for these organisms individually while maintaining the higher sensitivity of the two individual organisms (87%). In contrast, combining BVAB2 with *G. vaginalis* only very modestly improved sensitivity over either individual organism (84% to 85%), while specificity of the combination was lower than BVAB2 alone (91% compared with 93%). The combination of *A. vaginae* and BVAB2 improved sensitivity over either organism individually (92% compared with 79% to 87%), while specificity dropped slightly (90% compared with 91% to 93%). Finally, the combination of all three organisms provided better sensitivity (91%) than any individual organism or other combination, along with 93% specificity that was surpassed only by the combination of *G. vaginalis* and *A. vaginae* (94%). In general, these combinations significantly improved AUC in ROC analysis compared to individual organisms, with the exception of the combination of *G. vaginalis* and BVAB2, which was not superior to the use of *A. vaginae* alone (Table 4). When comparing the combination of all three organisms to combinations of two, superiority was observed versus the combinations of *G. vaginalis* and *A. vaginae* ($P = 0.039$) and versus *G. vaginalis* and BVAB2 ($P < 0.001$). We did not evaluate *Megasphaera* phylotypes 1 and 2 quantitatively, as they were present very rarely in non-BV specimens (0.0% and 1.9%, respectively). In addition, ROC analysis of *Lactobacillus* spp., either as individual species or in composite, revealed very low AUC values (≤0.46) due to their low prevalence among non-BV specimens (≤74%) (Table 2). Overall, these results support the concept that use of multiple organisms for the diagnosis of BV improves accuracy.

Generation and utilization of a logistic regression model for BV diagnosis. As we were working with a large number of variables that appeared to interact in a complex fashion with a categorical outcome (presence or absence of BV), we utilized logistic regression as a tool to model these interactions. Our first step was to determine which variables to include in our model. The utilization of a greater number of variables would increase the likelihood of the model; however, it also would increase the risk of overfitting, where the model describes error in the data set rather than actual relationships between variables and outcome. Therefore, we compared models incorporating all possible combina-

TABLE 4 *P* values for significance testing between the AUC of ROC curves (DeLong test)[a]

	P values by ROC curve for organism or model:						
AUC curve for organism or model	*G. vaginalis*	*A. vaginae*	BVAB2	*G. vaginalis* + *A. vaginae*	*G. vaginalis* + BV	*A. vaginae* + BV	*G. vaginalis* + *A. vaginae* + BV
A. vaginae	NS[b]						
BV	NS	0.025					
G. vaginalis + *A. vaginae*	<0.001	0.014	<0.001				
G. vaginalis + BVAB2	<0.001	NS	0.0017	0.0012			
A. vaginae + BVAB2	0.0032	0.0015	<0.001	NS	NS		
G. vaginalis + *A. vaginae* + BVAB2	<0.001	0.0046	<0.001	NS	<0.001	NS	
BIC[c] model	<0.001	<0.001	<0.001	0.0012	<0.001	0.0028	0.0048

[a] Only comparisons that were significant ($P < 0.05$) after controlling for multiple comparisons are reported.
[b] NS, nonsignificant.
[c] BIC, Bayesian information criterion.

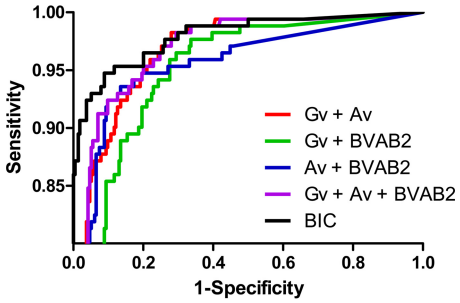

FIG 2 ROC analysis of qPCR data for combined BV-associated organisms and the Bayesian information criterion (BIC) logistic regression model. Av, *A. vaginae*; Gv, *G. vaginalis*.

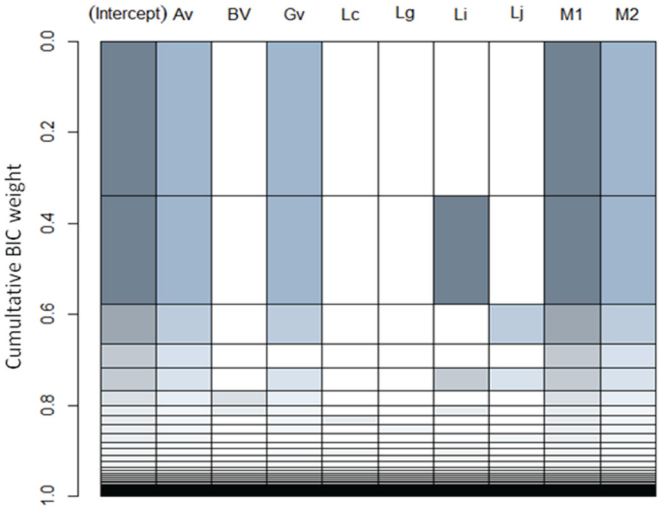

FIG 3 Plot of models ranked by Bayesian information criterion (BIC). Each horizontal band represents an individual model, and the height represents the portion of cumulative BIC weight across all models. Shading indicates the relative importance of each variable to an individual model. Abbreviations: Av, *A. vaginae*; BV, BVAB2; Gv, *G. vaginalis*; Lc, *L. crispatus*; Lg, *L. gasseri*; Li, *L. iners*; Lj, *L. jensenii*; M1, *Megasphaera* phylotype 1; M2, (*Megasphaera* phylotype 2.

tions of variables using the Bayesian information criterion (BIC), a maximum-likelihood criterion that is penalized for model complexity, defined as the incorporation of a greater number of variables. Thus, the BIC value is a function of the unexplained variance in the model and the number of variables included; therefore, models with lower BIC values are preferable to those with higher values. Using this method, we analyzed all 512 combinations of variables and found that the optimal model incorporated *A. vaginae*, *G. vaginalis*, *Megasphaera* phylotype 1, and *Megasphaera* phylotype 2 (Fig. 3). The relative importance of organisms in the models was as follows: *A. vaginae*, *Megasphaera* phylotype 1, and *Megasphaera* phylotype 2 (1.00); *G. vaginalis* (0.88); and *L. iners* (0.38), *L. jensenii* (0.21), BVAB2 (0.09), *L. crispatus* (0.06), and *L. gasseri* (0.05). As an independent method of variable selection, we also performed *k*-fold cross validation (CV). This method partitions the data into *k* subsamples, using a single subsample for testing the model and using the remaining subsamples as training data. We utilized 10 partitions, a standard number for this phylotype of analysis (i.e., 10-fold CV). This method identified the same logistic regression model as the best possible model (data not shown). ROC analysis (Fig. 2) clearly showed that the BIC-selected model was an improvement over our prior analyses (Fig. 1). Although its sensitivity (92%) was similar to models using simple additive titers of organisms, its specificity was higher than these models at 95% (Table 3). The BIC-selected model had a significantly greater AUC than every other model we tested (Table 4) and correctly classified 157 of 171 BV specimens, 204 of 214 non-BV specimens, and 361 of 385 total specimens, resulting in a diagnostic test with 94% accuracy, 92% sensitivity, 95% specificity, 94% positive predictive value (PPV), and 94% negative predictive value (NPV) (Table 5). Therefore, the use of logistic regression modeling allowed us to improve the accuracy of BV diagnosis using PCR data.

One important consideration is that most of the participants in our study were sampled multiple times, possibly confounding our analysis. To address this issue, we tested the accuracy of the BIC model on specimens obtained at each of the first 3 subject visits, in which each specimen represented a single individual (Table 5). Sensitivity and specificity at each of these visits was ≥91%, indicating that the model is robust in several clinical settings. For example, when we analyzed just the first visit, when most samples (84%) were from BV-positive women, our model had a sensitivity of 91% and a specificity of 96%. At the second visit, when most women (94%) were BV negative, the sensitivity and specificity were 100% and 96%, respectively. At the third visit, when moderate proportion of women had BV (38%), the sensitivity and specificity were both 91%. Overall, these results indicate that multiple sampling from study participants did not substantially impact our findings. Last, we evaluated the accuracy of the BIC-selected model on women with intermediate flora as opposed to normal flora (for women without BV) or abnormal flora (for women with BV) (Table 6). With regard to women without BV and normal flora, our model correctly classified 171 of 174 specimens (98%); for women with BV and abnormal flora, the model correctly classified 144 of 151 specimens (95%). Our model was somewhat less accurate in identification of specimens from women with intermediate flora, correctly classifying 33 of 40 specimens (83%) from women without BV and 13 of 20 (65%) from women with BV.

DISCUSSION

In this report, we discuss the development of a highly accurate qPCR assay for BV diagnosis. This study verifies and expands upon a number of prior studies evaluating PCR-based assays for BV diagnosis. Fredricks et al. (17) assayed 17 different organisms by conventional PCR and evaluated their sensitivity and specificity in comparison to Amsel's criteria and Nugent score. The highest accuracy was reported for detection of BVAB2 or *Megasphasera* phylotype 1, with sensitivity and specificity of 99% and 89%, respectively, compared to Amsel's criteria and 96% and 99%, respectively, compared to Nugent scoring. A subsequent study using qPCR found that the presence of $>10^9$ *G. vaginalis* or $>10^8$ *A. vaginae* genomic copies in patient specimens was 100% sensitive and 93% specific for BV diagnosis (18). Although both of these assays are highly accurate, it is important to note that they only evaluated specimens from subjects with Nugent scores of ≤3 or ≥7; patients with intermediate flora (Nugent score of 4 to 6) were not included in either analysis. However, clinical BV by Amsel's criteria can be present in individuals with intermediate Nugent scores as well as in those with abnormal flora. Prior studies

TABLE 5 Performance of BIC-selected model on specimens by visit

Subgroup or performance evaluation	Value by visit: Initial visit	Post-treatment (1 week)	Short-term follow-up (40–45 Days)	Long-term follow-up (~monthly)	Total
No. of women	146	103	85	24	148
No. of samples	146	103	85	51	385
No. (%) with Amsel's criteria 0–2 (non-BV)	23	97	53	41	214
No. (%) correct[a]	22 (96)	93 (96)	48 (91)	41 (100)	204 (95)
No. (%) incorrect[a]	1 (4.3)	4 (4.1)	5 (9.4)	0 (0.0)	10 (4.7)
No. (%) with Amsel's criteria 3–4 (BV)	123 (84)	6	32	10	171
No. (%) correct[b]	112 (91)	6 (100)	29 (91)	10 (100)	157 (92)
No. (%) incorrect[b]	11 (8.9)	0 (0.0)	3 (9.4)	0 (0.0)	14 (8.2)
Sensitivity, % (95% CI[c])	91 (85–95)	100 (54–100)	91 (75–98)	100 (69–100)	92 (87–95)
Specificity, % (95% CI)	96 (78–100)	96 (90–99)	91 (79–97)	100 (91–100)	95 (92–98)
Positive predictive value, % (95% CI)	99 (95–100)	60 (26–88)	85 (69–95)	100 (96–100)	94 (89–97)
Negative predictive value, % (95% CI)	67 (48–82)	100 (96–100)	94 (84–99)	100 (91–100)	94 (89–96)

[a] Percentage is a percent of the specimens with 0–2 Amsel's criteria.
[b] Percentage is a percent of the specimens with 3–4 Amsel's criteria.
[c] CI, confidence interval.

found that 37% to 54% of subjects with intermediate flora by Gram stain had BV by Amsel's criteria (26, 27). Overall, specimens from subjects with intermediate flora represented 16% (60 of 385) of the total specimens included in the present study. Specifically, we found that 20 of 60 (33%) subjects with intermediate flora had BV by Amsel's criteria, accounting for 12% (20 of 171) of BV subjects. In general, categorization of these subjects by our PCR results was less accurate than for women with either normal or abnormal flora, which may partially account for their exclusion from prior studies. In our study, intermediate flora specimens represented 14 of 24 (58%) incorrectly classified specimens, despite only accounting for 16% of specimens analyzed. Indeed, if we excluded these specimens, our accuracy would increase from 94% to 97%; sensitivity, from 92% to 95%; specificity, from 95% to 98%; PPV, from 94% to 98%; and NPV, from 94% to 96%. However, as any diagnostic tool will realistically have to categorize specimens with intermediate flora as being BV or non-BV for purposes of patient treatment, we propose that our accuracy when including intermediate specimens, although lower, is a more clinically relevant representation of this assay.

Last, in a study that evaluated the use of qPCR assays quantifying *G. vaginalis*, *A. vaginae*, BVAB2, *Megasphaera* phylotype 1, and *L. crispatus* organisms and that utilized logistic regression to select organisms for inclusion, the investigators found that detection of *A. vaginae*, *Megasphaera* phylotype 1, and BVAB2 provided the most accurate assay, with 97% sensitivity and 92% specificity (33). Although this study did include intermediate flora specimens, it still categorized 5.3% of samples as indeterminate after testing, thus limiting its clinical utility. Our study differs from that of Cartwright et al. (33) in that we (i) included a greater number of potential organisms (9 versus 5), (ii) avoided the use of an indeterminate category for ambiguous specimens but rather designated all nonexcluded specimens as either BV or non-BV, as the decision of the clinician to treat BV is similarly binary, and (iii) utilized BIC and CV for selection of logistic regression models to avoid overfitting. Interestingly, despite a similar approach in both studies, our study came to a different conclusion with respect to which organisms provided the best diagnostic value. Both studies identified *A. vaginae* and *Megasphaera* phylotype 1 as indicator organisms; however, the study by Cartwright et al. utilized BVAB2, and ours incorporated *G. vaginalis* (33). Similar to Menard et al. (18), we found that *G. vaginalis* can be a highly accurate indicator of BV if quantitation and ROC analysis are performed. It is noteworthy that Cartwright et al. (33) found that the specificity of BVAB2 detection (present in 11 of 61 BV-negative specimens) for BV diagnosis was 82%, the same as our study, which was lower than what was observed by Fredricks et al. (93% to 97%) (17). This result is most likely due to the use of more sensitive qPCR methods by Cartwright et al. and less sensitive conventional PCR by the Fredricks et al. Consistent with this reasoning, the utilization of ROC analysis allowed us to improve specificity to 93%, as the concentration of BVAB2 in BV-negative specimens is generally very low. In contrast to Cartwright et al., our study also considered *Megasphaera* phylotype 2 for inclusion in a diagnostic panel. Although its low sensitivity for detection of BV-positive specimens (7% to 19%) by conventional PCR (17) would appear to make it a

TABLE 6 Accuracy of BIC model for BV diagnosis in specimens stratified by Nugent score

Specimen group	Nugent score	BIC model accuracy (%) of BV vs non-BV diagnosis		
Non-BV (Amsel's 0–2)	Normal (0–3)	3	171	98
	Intermediate (4–6)	7	33	83
BV (Amsel's 3–4)	Intermediate (4–6)	13	7	65
	Abnormal (7–10)	144	7	95

poor candidate, we found that, by using qPCR, *Megasphaera* phylotype 2 was 31% sensitive and 100% specific for BV diagnosis and, indeed, was included in our optimized models. In addition, both the study by Cartwright et al. and our study considered *L. crispatus*, an organism associated with vaginal health, as a potential negative indicator of BV; in both cases, though, the organism was not found to contribute sufficiently to be included in the final panel of organisms for detecting BV by PCR and ROC analysis. In addition, we evaluated two other similar species, *L. gasseri* and *L. jensenii*, and they did not merit inclusion either. In total, we found that the absence of these three organisms is only 70% sensitive and 74% specific for BV, resulting in accuracy inferior to that of individual BV organisms, such as *Megasphaera* phylotype 1 or BVAB2. Overall, the prevalence of *Lactobacillus* spp. (other than *L. iners*) was rather low in our study, which may be due to the fact that 68% of our study participants were black, a population in which colonization with these species is less frequently observed than in Caucasian women (32). In fact, the prevalence of *L. crispatus* in our study was similar to that in the study by Cartwright et al. (33), at 41% in non-BV specimens and 18% in BV specimens in our study compared to 43% in non-BV specimens and 11% in BV specimens in their study, in which 87% of the participants were black. Although these *Lactobacillus* spp. were not included in optimal assays for initial diagnosis (as defined by CV and/or BIC values), that does not mean that they do not provide diagnostic value. For example, recovery of the *Lactobacillus* spp. after BV treatment may very well be a useful prognostic marker for long-term sustained response to treatment, and conversely, their absence may identify patients at high risk for recurrent disease. In addition, as BV probiotic therapies continue to evolve (34), it will be useful to test for the presence of *Lactobacillus* spp. to identify patients who will benefit from such therapies and to monitor these patients posttreatment to determine if colonization with these organisms is transient or long-term. Additionally, further analysis of the diagnostic utility, and combination, of qPCR assays for BV-associated bacteria and *Lactobacillus* spp. in a relative concentration evaluation model to accurately differentiate not only symptomatic BV samples but also asymptomatic BV samples, symptomatic and asymptomatic intermediate samples, and transitional vaginal samples is merited. Last, we evaluated *L. iners*, but this organism is known to be present in individual with BV and without BV (21), so it was not surprising to see that it offered little diagnostic value for determining BV flora. However, tracking this organism through patient treatment may be useful (35).

In summary, we evaluated a broad variety of vaginal organisms as potential diagnostic markers of BV using a large collection of well-characterized specimens, even, including those from subjects with intermediate flora, which are frequently excluded from diagnostic BV studies despite their significant prevalence. We found that quantitation and ROC analysis greatly improved the diagnostic values, as did combinations of BV-associated organisms compared to individual ones. We used logistic regression analysis and model selection to identify an optimized set of organisms for diagnostic testing, which significantly improved their accuracy while preventing overfitting by using penalized maximum likelihood criteria. This diagnostic tool is highly accurate and useful for both clinical diagnostic and research purposes.

FUNDING INFORMATION

All funding was provided by the Genesis Biotechnology group.

REFERENCES

1. **Koumans EH, Sternberg M, Bruce C, McQuillan G, Kendrick J, Sutton M, Markowitz LE.** 2007. The prevalence of bacterial vaginosis in the United States, 2001-2004: associations with symptoms, sexual behaviors, and reproductive health. Sex Transm Dis **34:**864–869. http://dx.doi.org/10.1097/OLQ.0b013e318074e565.
2. **Sobel JD.** 2000. Bacterial vaginosis. Annu Rev Med **51:**349–356. http://dx.doi.org/10.1146/annurev.med.51.1.349.
3. **Amsel R, Totten PA, Spiegel CA, Chen KC, Eschenbach D, Holmes KK.** 1983. Nonspecific vaginitis: diagnostic criteria and microbial and epidemiologic associations. Am J Med **74:**14–22. http://dx.doi.org/10.1016/0002-9343(83)91112-9.
4. **Leitich H, Kiss H.** 2007. Asymptomatic bacterial vaginosis and intermediate flora as risk factors for adverse pregnancy outcome. Best Pract Res Clin Obstet Gynaecol **21:**375–390. http://dx.doi.org/10.1016/j.bpobgyn.2006.12.005.
5. **Martin HL, Richardson BA, Nyange PM, Lavreys L, Hillier SL, Chohan B, Mandaliya K, Ndinya-Achola JO, Bwayo J, Kreiss J.** 1999. Vaginal lactobacilli, microbial flora, and risk of human immunodeficiency virus type 1 and sexually transmitted disease acquisition. J Infect Dis **180:**1863–1868. http://dx.doi.org/10.1086/315127.
6. **Schwebke JR, Desmond R.** 2007. A randomized trial of metronidazole in asymptomatic bacterial vaginosis to prevent the acquisition of sexually transmitted diseases. Am J Obstet Gynecol **196:**517.e1–517.e6. http://dx.doi.org/10.1016/j.ajog.2007.02.048.
7. **Hillier SL, Nugent RP, Eschenbach DA, Krohn MA, Gibbs RS, Martin DH, Cotch MF, Edelman R, Pastorek JG, II, Rao AV, McNellis D, Regan JA, Carey JC, Klebanoff MA.** 1995. Association between bacterial vaginosis and preterm delivery of a low-birth-weight infant: the vaginal infections and prematurity study group. N Engl J Med **333:**1737–1742. http://dx.doi.org/10.1056/NEJM199512283332604.
8. **Persson E, Bergstrom M, Larsson PG, Moberg P, Platz-Christensen JJ, Schedvins K, Wolner-Hanssen P.** 1996. Infections after hysterectomy: a prospective nation-wide Swedish study—the study group on Infectious Disease in Obstetrics and Gynecology within the Swedish Society of Obstetrics and Gynecology. Acta Obstet Gynecol Scand **75:**757–761. http://dx.doi.org/10.3109/00016349609065742.
9. **Larsson PG, Platz-Christensen JJ, Thejls H, Forsum U, Pahlson C.** 1992. Incidence of pelvic inflammatory disease after first-trimester legal abortion in women with bacterial vaginosis after treatment with metronidazole: a double-blind, randomized study. Am J Obstet Gynecol **166:**100–103. http://dx.doi.org/10.1016/0002-9378(92)91838-2.
10. **Balkus JE, Richardson BA, Rabe LK, Taha TE, Mgodi N, Kasaro MP, Ramjee G, Hoffman IF, Abdool Karim SS.** 2014. Bacterial vaginosis and the risk of trichomonas vaginalis acquisition among HIV-1-negative women. Sex Transm Dis **41:**123–128. http://dx.doi.org/10.1097/OLQ.0000000000000075.
11. **Landers DV, Wiesenfeld HC, Heine RP, Krohn MA, Hillier SL.** 2004. Predictive value of the clinical diagnosis of lower genital tract infection in women. Am J Obstet Gynecol **190:**1004–1010. http://dx.doi.org/10.1016/j.ajog.2004.02.015.
12. **Nugent RP, Krohn MA, Hillier SL.** 1991. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. J Clin Microbiol **29:**297–301.
13. **Gardner HL, Dukes CD.** 1955. Haemophilus vaginalis vaginitis: a newly defined specific infection previously classified nonspecific vaginitis. Am J Obstet Gynecol **69:**962–976.
14. **Spiegel CA, Amsel R, Eschenbach D, Schoenknecht F, Holmes KK.** 1980. Anaerobic bacteria in nonspecific vaginitis. N Engl J Med **303:**601–607. http://dx.doi.org/10.1056/NEJM198009113031102.
15. **Ferris MJ, Masztal A, Aldridge KE, Fortenberry JD, Fidel PL, Jr, Martin DH.** 2004. Association of Atopobium vaginae, a recently described metronidazole resistant anaerobe, with bacterial vaginosis. BMC Infect Dis **4:**5. http://dx.doi.org/10.1186/1471-2334-4-5.
16. **Fredricks DN, Fiedler TL, Marrazzo JM.** 2005. Molecular identification of bacteria associated with bacterial vaginosis. N Engl J Med **353:**1899–1911. http://dx.doi.org/10.1056/NEJMoa043802.
17. **Fredricks DN, Fiedler TL, Thomas KK, Oakley BB, Marrazzo JM.** 2007. Targeted PCR for detection of vaginal bacteria associated with bacterial vaginosis. J Clin Microbiol **45:**3270–3276. http://dx.doi.org/10.1128/JCM.01272-07.
18. **Menard JP, Fenollar F, Henry M, Bretelle F, Raoult D.** 2008. Molecular

quantification of Gardnerella vaginalis and Atopobium vaginae loads to predict bacterial vaginosis. Clin Infect Dis **47:**33–43. http://dx.doi.org/10.1086/588661.

19. **Balashov SV, Mordechai E, Adelson ME, Sobel JD, Gygax SE.** 2014. Multiplex quantitative polymerase chain reaction assay for the identification and quantitation of major vaginal lactobacilli. Diagn Microbiol Infect Dis **78:**321–327. http://dx.doi.org/10.1016/j.diagmicrobio.2013.08.004.
20. **Ma B, Forney LJ, Ravel J.** 2012. Vaginal microbiome: rethinking health and disease. Annu Rev Microbiol **66:**371–389. http://dx.doi.org/10.1146/annurev-micro-092611-150157.
21. **Zozaya-Hinchliffe M, Lillis R, Martin DH, Ferris MJ.** 2010. Quantitative PCR assessments of bacterial species in women with and without bacterial vaginosis. J Clin Microbiol **48:**1812–1819. http://dx.doi.org/10.1128/JCM.00851-09.
22. **Klebanoff MA, Schwebke JR, Zhang J, Nansel TR, Yu KF, Andrews WW.** 2004. Vulvovaginal symptoms in women with bacterial vaginosis. Obstet Gynecol **104:**267–272. http://dx.doi.org/10.1097/01.AOG.0000134783.98382.b0.
23. **Chernesky M, Castriciano S, Jang D, Smieja M.** 2006. Use of flocked swabs and a universal transport medium to enhance molecular detection of Chlamydia trachomatis and Neisseria gonorrhoeae. J Clin Microbiol **44:**1084–1086. http://dx.doi.org/10.1128/JCM.44.3.1084-1086.2006.
24. **Balashov SV, Mordechai E, Adelson ME, Gygax SE.** 2014. Identification, quantification and subtyping of Gardnerella vaginalis in noncultured clinical vaginal samples by quantitative PCR. J Med Microbiol **63:**162–175. http://dx.doi.org/10.1099/jmm.0.066407-0.
25. **Smith WL, Hedges SR, Mordechai E, Adelson ME, Trama JP, Gygax SE, Kaunitz AM, Hilbert DW.** 2014. Cervical and vaginal flora specimens are highly concordant with respect to bacterial vaginosis-associated organisms and commensal lactobacillus species in women of reproductive age. J Clin Microbiol **52:**3078–3081. http://dx.doi.org/10.1128/JCM.00795-14.
26. **Taylor-Robinson D, Morgan DJ, Sheehan M, Rosenstein IJ, Lamont RF.** 2003. Relation between Gram-stain and clinical criteria for diagnosing bacterial vaginosis with special reference to Gram grade II evaluation. Int J STD AIDS **14:**6–10. http://dx.doi.org/10.1258/095646203321043183.
27. **Bradshaw CS, Morton AN, Garland SM, Horvath LB, Kuzevska I, Fairley CK.** 2005. Evaluation of a point-of-care test, BVBlue, and clinical and laboratory criteria for diagnosis of bacterial vaginosis. J Clin Microbiol **43:**1304–1308. http://dx.doi.org/10.1128/JCM.43.3.1304-1308.2005.
28. **Workowski KA, Bolan GA, Centers for Disease Control and Prevention.** 2015. Sexually transmitted diseases treatment guidelines, 2015. MMWR Morb Mortal Wkly Rep **64:**1–137.
29. **Pheifer TA, Forsyth PS, Durfee MA, Pollock HM, Holmes KK.** 1978. Nonspecific vaginitis: role of Haemophilus vaginalis and treatment with metronidazole. N Engl J Med **298:**1429–1434. http://dx.doi.org/10.1056/NEJM197806292982601.
30. **Verstraelen H, Verhelst R, Claeys G, Temmerman M, Vaneechoutte M.** 2004. Culture-independent analysis of vaginal microflora: the unrecognized association of Atopobium vaginae with bacterial vaginosis. Am J Obstet Gynecol **191:**1130–1132. http://dx.doi.org/10.1016/j.ajog.2004.04.013.
31. **Antonio MA, Hawes SE, Hillier SL.** 1999. The identification of vaginal Lactobacillus species and the demographic and microbiologic characteristics of women colonized by these species. J Infect Dis **180:**1950–1956. http://dx.doi.org/10.1086/315109.
32. **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, Forney LJ.** 2011. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci U S A **108**(suppl 1)**:**S4680–S4687. http://dx.doi.org/10.1073/pnas.1002611107.
33. **Cartwright CP, Lembke BD, Ramachandran K, Body BA, Nye MB, Rivers CA, Schwebke JR.** 2012. Development and validation of a semi-quantitative, multitarget PCR assay for diagnosis of bacterial vaginosis. J Clin Microbiol **50:**2321–2329. http://dx.doi.org/10.1128/JCM.00506-12.
34. **Huang H, Song L, Zhao W.** 2014. Effects of probiotics for the treatment of bacterial vaginosis in adult women: a meta-analysis of randomized clinical trials. Arch Gynecol Obstet **289:**1225–1234. http://dx.doi.org/10.1007/s00404-013-3117-0.
35. **Verstraelen H, Verhelst R, Claeys G, De Backer E, Temmerman M, Vaneechoutte M.** 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. http://dx.doi.org/10.1186/1471-2180-9-116.