1 Diagnosing Bacterial Vaginosis with a Novel, Clinically-Actionable Molecular Diagnostic Tool

2

3 Joseph P. Jarvis,[a] Doug Rains,[b] Steven J. Kradel,[a] James Elliott,[b] Evan E. Diamond,[c] Erik

4 Avaniss-Aghajani,[d] Farid Yasharpour,[e] Jeffrey A. Shaman[a#]

5

6 [a]Coriell Life Sciences, Philadelphia, Pennsylvania, USA

7 [b]Quantigen Genomics, Fishers, Indiana, USA

8 [c]ThermoFisher Scientific

9 [d]Primex Clinical Laboratories, Van Nuys, California, USA

10 [e]Maternity & Infertility Institute, San Fernando, California, USA

11

12 Running Head: Novel Molecular Diagnostic for Bacterial Vaginosis

13

14 #Address correspondence to Jeffrey A. Shaman, jshaman@coriell.com.

Abstract/Summary

Bacterial vaginosis is a common condition among women of reproductive age and is associated with potentially serious side-effects, including an increased risk of preterm birth. Recent advancements in microbiome sequencing technologies have produced novel insights into the complicated mechanisms underlying bacterial vaginosis and have given rise to new methods of diagnosis. Here we report on the validation of a quantitative, molecular diagnostic algorithm based on the relative abundances of ten potentially pathogenic bacteria and four commensal Lactobacillus species in research subjects (n = 172) classified as symptomatic (n = 149) or asymptomatic (n = 23). We observe a clear and reinforcing pattern among patients diagnosed by the algorithm that is consistent with the current understanding of biological dynamics and dysregulation of the vaginal microbiome during infection. Using this enhanced assessment of the underlying biology of infection, we demonstrate improved diagnostic sensitivity (93%) and specificity (90%) relative to current diagnostic tools. Our algorithm also appears to provide enhanced diagnostic capabilities in ambiguous classes of patients for whom diagnosis and medical decision-making is complicated, including asymptomatic patients and those deemed "intermediate" by Nugent scoring. Ultimately, we establish CLS2.0q as a quantitative, sensitive, specific, accurate, robust, and flexible algorithm for the clinical diagnosis of bacterial vaginosis – importantly, one that is also ideal for the differential diagnosis of non-BV infections with clinically similar presentations.

Introduction

36 Bacterial vaginosis (BV) is widely considered to be the most common vaginal disorder
37 among women of reproductive age (1). In a sample of self-collected vaginal swabs from 3,700
38 women, the prevalence of BV was estimated at 29% in the general population of women aged
39 14-49 years and 50% in African-American women (2). Typically, affected patients present with
40 inflammation of the vagina that can result in discharge, itching, and pain. That said, clinical
41 presentation can vary considerably, with asymptomatic infections being noted with great
42 regularity (2). A wide variety of risk factors for BV have been noted, including sexual activity
43 (3,4), the presence of other sexually transmitted infections (5), as well as race and ethnicity (2).
44 Important complications of BV infection include an increased risk of preterm delivery in
45 pregnant women (6) and of HIV acquisition and transmission (7). In some women, clearance of
46 the infection, even with standard of care, is difficult to achieve, often resulting in chronic or
47 recurrent symptoms (8).
48 Improved strategies for detecting BV infection have been the subject of much discussion,
49 due to the reduction in the quality of life that can result from an active infection and the
50 reasonably high prevalence of BV in women (9). Recent advances in technology have driven
51 interest in DNA-based, quantitative, diagnostic algorithms that are theoretically quicker to
52 perform than culture, more accurate in their results, and are potential candidates as point-of-care
53 devices. However, the seemingly complex biology and dynamics of BV infection, along with
54 important limitations in "gold-standard" approaches, have complicated these efforts.
55 One confounding factor is the consistent observation that BV is associated not with a
56 single causative pathogen, but rather with a broad disruptive shift in the microbial population:
57 from a healthy *Lactobacilli*-dominated environment to one largely overtaken by organisms such
58 as *Gardnerella,* various mycoplasma, and a heterogeneous collection of facultative anaerobes

59 (10). With this shift come changes in the underlying interactions among different organisms,
60 changes in the metabolites produced by the local bacterial community, and changes to the host's
61 immune response. Together, these factors produce key signs and symptoms of infection. This
62 dynamic mechanism of disease manifestation results in a condition with all the standard
63 hallmarks of a complex etiology. For example, BV presents clinically in numerous ways, and
64 individual symptoms can be the result of multiple systematic shifts in the vaginal microbiome in
65 favor of a variety of pathogenic organisms. Sometimes women experience no clinical
66 symptoms; in fact, a significant percentage of affected patients are asymptomatic (11,33). This
67 statistic is particularly important in the context of pregnancy, where it has been shown that BV
68 infection, whether symptomatic or not, is associated with preterm birth (6).

69 In addition, the vaginal microbiota has recently been shown to be longitudinally dynamic
70 in healthy individuals; that is, changes are seen to occur in near-real time over days and weeks
71 (12). The impact of microbial population flux on the maintenance of bacterial community
72 homeostasis, on the resolution of symptoms post-infection, and on the repression of symptoms in
73 asymptomatic women is not well understood. Nevertheless, an oft-changing microbiota poses
74 difficulties for accurate diagnosis, since efficient molecular testing depends on samples taken at a
75 singular point in time.

76 Given the complexity of the dynamics that lead to clinical presentation, it is unsurprising
77 that diagnosis of BV infections has historically been a challenge. While Gram staining plus
78 Nugent scoring (13) is considered the gold-standard, it is a technique fraught with problems: 1.
79 Nugent requires expertise at staining and scoring morphotypes; 2. several clinically significant
80 microbes, either of variable morphology or lacking a cell wall, are regularly misidentified or
81 overlooked (9); and 3. some 20% of symptomatic subjects score between 4 and 6, rendering

82 their status “intermediate”—an ambiguous diagnosis that is not helpful for the healthcare
83 provider (14) and may lead to inappropriate clinical intervention, or, conversely, to no
84 intervention when it is actually needed. Given the complexity of microbiome alterations now
85 known to be associated with disease, the second of these shortcomings may account for Nugent’s
86 reportedly poor sensitivity and specificity (13). Finally, Nugent scoring has also been shown to
87 occasionally miss differential diagnoses such as *Trichomonas vaginalis* infections (15) with
88 potentially problematic effects on patient well-being due to inappropriate therapy.

89 An older and more commonly used diagnostic tool is the application of Amsel’s criteria.
90 This test, first described in 1983, requires the satisfaction of at least three of four criteria to
91 qualify for a BV diagnosis: elevated pH (>4.5), characteristic discharge, fishy odor upon the
92 addition of 10% KOH to a slide prep, and the microscopic visualization of “clue cells” (smaller
93 bacteria adhering to epithelial cells) in a wet mount (16). However, while easy to implement in
94 theory, Amsel’s criteria are reportedly underutilized in clinical settings (17) and are technically
95 limited by the lack of consideration of the abundance of healthy *Lactobacillus* species (18).

96 These and other acknowledged failings of the gold-standard diagnostic techniques for BV
97 have recently fueled a search for alternatives leveraging new technology. Today, there are a
98 variety of alternatives in the marketplace that claim ease of testing, enhanced sensitivity and
99 specificity, as well as reduced cost. Relatively few of these can be used at the point-of care
100 (PoC) (19,20); rather, most available tests require sending a sample – often in the form of a
101 vaginal swab – to a laboratory for testing. Though not as expedient as a PoC diagnostic, at 1 to 2
102 days for a result after sample transport and testing, versus minutes to hours for a PoC result, the
103 referencing of samples to an external laboratory offers several key advantages. Most notably,
104 some clinical laboratories can simultaneously screen samples for a broad range of microbial

105 targets covering multiple conditions, not merely BV. The ability to offer a differential diagnosis
106 is a key advantage for symptomatic patients, since infections other than BV can present with
107 very similar signs and symptoms. Many labs can also generate bacteria-specific
108 presence/absence or semi-quantitative information, while also detecting difficult-to-culture,
109 anaerobic, and commensal microorganisms. Such data offer the potential for healthcare
110 providers, who are often faced with interpreting ambiguous signs and symptoms, to receive
111 greater diagnostic clarity. Similarly, they can be alerted to mixed/co-infections, a situation
112 reported to affect 20% of BV+ patients (21). Finally, providers can potentially receive specific
113 recommendations for therapy (e.g. antibiotic regimens or combinations of treatment options
114 specific to the pathogen(s) detected). While in no way a replacement for a thorough physical
115 examination and detailed capture of patient history (22), such molecular-based screening holds
116 great promise for the enhancement of the current standard-of-care.
117 Here we present the refinement and performance evaluation of an existing quantitative
118 molecular diagnostic for BV that eliminates the challenges demonstrated by Amsel and Nugent
119 tests, while also providing an accurate diagnosis via genus and species-specific information for
120 both pathogenic and non-pathogenic bacteria. The test, based on real-time PCR, is offered in a
121 highly customizable format, allowing individual laboratories to add or subtract content as future
122 needs and research developments dictate. The approach also permits simultaneous screening of
123 an assortment of other vaginal pathogens with no additional labor or consumable requirements,
124 thereby enabling differential diagnoses – an added value of this test over current PoC solutions.
125

126 # Materials and Methods

127 *Diagnostic Panel Design*

128 As part of previous BV diagnostic algorithm development efforts, we conducted a
129 comprehensive review of published studies on normal and pathogenic vaginal microbiota. Our
130 survey included literature characterizing the microbiomes of BV+ and BV- women via next-
131 generation sequencing (12,17,23,24), a survey of molecular-based diagnostic tool development
132 for BV (1,9,14,25,26,27), and expert reviews of both of these subjects (28,29,30,31,32). We
133 found that numerous studies have shown that the amounts of particular bacteria – and not merely
134 their presence or absence – play a role in the development and/or persistence of BV
135 (1,9,14,17,27,28,31,34,35). Ultimately, fourteen BV-associated microorganisms (TABLE 1) that
136 would theoretically maximize diagnostic specificity while remaining cost-effective to the clinical
137 laboratory were selected for inclusion in the new diagnostic test. TaqMan®-based assays were
138 designed with the assistance of ThermoFisher's bioinformatic pipeline to detect bacterial-specific
139 genes down to the species level, reported to be present at one copy per microbe, a fact important
140 for the relative abundance calculation of each microbial target. The assays were subsequently
141 spotted into the through-holes of a custom-designed TaqMan OpenArray plate for vaginal
142 microbiota investigations (available from ThermoFisher Scientific).

143 Since bacterial vaginosis, candidiasis, and trichomoniasis are all indicated in vaginitis,
144 present clinically with similar signs and symptoms, and are often observed in co-infected
145 patients, assays specific to these infections were added to the panel as an aid to differential
146 diagnoses. Furthermore, because *Candida albicans* infections are treated differently than non-
147 *albicans* infections, five specific yeast species were also included in the final panel design.

148 *Algorithm Development and Study Population*

149 During the original algorithm design phase, several statistical and algorithmic approaches
150 were applied to a sample of patients (410 antibiotic-naïve, including 25 asymptomatic women

151 showing no clue cells on analysis for Amsel's criteria) and evaluated with respect to sensitivity
152 and specificity. These included a number of multivariate linear and mixed models as well as a
153 variety of other techniques. In general, linear regression techniques performed poorly with
154 respect to sensitivity (i.e., the ability to identify BV+ individuals), but performed fairly well with
155 respect to specificity (i.e., the ability to exclude BV as a diagnosis). Alternative strategies, such
156 as those that produced the final CLS1 algorithm, showed improved overall performance
157 (especially increased sensitivity), albeit with slightly reduced specificity. Ultimately, the highest
158 performing quantitative model using measures of relative abundance was chosen for validation,
159 as opposed to one based solely on presence/absence data.

160 Subsequently, a total of 172 research subjects (149 symptomatic and 23 asymptomatic
161 women) were recruited into a clinical study aimed at validating the existing, proprietary
162 diagnostic algorithm (CLS2.0q). All were enrolled at the Maternity & Infertility Institute (San
163 Fernando, CA). Patients younger than 18 years old or those who had taken antibiotics in the
164 previous 30 days were excluded. Patient ethnicity, age, pregnancy status, menstruation status,
165 history of recurrent bacterial vaginosis and candidiasis, and method of birth control were
166 recorded. Nugent scoring was also performed on each. The average age of enrollees was 30.4
167 years; 91.8% self-classified as Hispanic/Latina.

168 *Sample collection and analysis*

169 As in the population used for algorithm development, three samples from the lateral
170 vaginal wall and posterior fornix or a blind vaginal swab were collected in the validation
171 population using COPAN ESwab™ with flocked nylon fiber tips (36). Two of the swabs were
172 transferred into 1 mL of modified Liquid Amies Transport medium and shipped at room

173 temperature to the laboratories for DNA processing. One swab was used to transfer collected
174 sample material onto each of three slides.

175 One slide was examined for *Trichomonas vaginalis* and clue cells using standard
176 microscopy techniques. The second slide was examined using the whiff test by adding a drop of
177 10% KOH solution and checking for a characteristically strong fishy odor. The third slide was
178 sent to Primex Clinical Laboratories (Van Nuys, CA) for independent evaluation by a pathologist
179 and technician blinded to the clinical presentation of the sample's donor.

180 The pathologist and technician applied Nugent scoring methods (13) to a smear of a
181 sample collected from each patient. Briefly, the swab was rolled on a glass slide; the smear was
182 then heat-fixed, and Gram stained. The smear was subsequently evaluated for the following
183 morphotypes (1000x magnification under oil immersion): large Gram-positive rods
184 (*Lactobacillus* morphotypes), small Gram-variable rods (*G. vaginalis* morphotypes), small
185 Gram-negative rods (*Bacteroides* species morphotypes), and curved Gram-variable rods
186 (*Mobiluncus* species morphotypes). Nugent scoring was used to diagnose bacterial vaginosis.
187 For samples with a 4-6 "intermediate" score, the presence of three of four the Amsel's criteria
188 (grayish-white homogeneous vaginal discharge, pH≥4.5, positive whiff test, and presence of clue
189 cells) was used as the criterion to diagnose BV.

190 For the molecular assays, DNA was isolated using MagMAX™ Express 96 semi-
191 automated sample prep system in combination with the MagMAX™ DNA Multi-Sample Ultra
192 Kit (ThermoFisher Scientific, Waltham, MA) and standard protocols. Next, real-time PCR was
193 performed on multiple replicates using 2.5 uL of 2X OpenArray® Gene Expression Master Mix
194 combined with 2.5 uL eluted DNA in a 384-well plate and added to the OpenArray plate using
195 the AccuFill™ (ThermoFisher Sc.) automated pipetting instrument. OpenArray plates were

196 cycled on the QuantStudio 12K Flex OpenArray Real-Time PCR System (ThermoFisher Sc.)
197 according to the manufacturer's recommended conditions. Following cycling, data were
198 analyzed and exported using QuantStudio software. Samples were only counted as true positives
199 when at least 50% of technical replicates for a given microbe-specific Assay yielded a Crt value
200 ≤31. Additionally, any sample in which the 16S rDNA control failed to achieve a minimal
201 threshold of 21 was rejected, since higher values suggest inadequate sample collection. Relative
202 quantities for each BV-associated bacterial assay were calculated based on Crt values for each of
203 the targets.
204

205 # Results

206 Clinical samples from 149 symptomatic and 23 asymptomatic women were tested for the
207 relative abundance of the 14 distinct bacterial species (TABLE 1) that were chosen to provide a
208 clinically-significant representation of the vaginal microbiome. Based on these data, patients
209 were diagnosed as either BV+ or BV- using a previously developed, proprietary, quantitative
210 molecular diagnostic algorithm, CLS2.0q (FIG 1a). Sensitivity and specificity were assessed
211 with respect to the gold-standard Nugent test and found to be 93% and 90%, respectively.
212 Several clear patterns emerged when comparing diagnostic outcomes with observed
213 variation in bacterial abundance. First, BV+ women were found to have relatively high levels of
214 multiple pathogenic organisms. No single organism perfectly differentiated between BV+ and
215 BV- patients; rather, the combined presence of several microbes (especially Megasphaera 1,
216 BVAB2, *A. vaginae*, and *G. vaginalis*) contribute to a clear signal of disease. Furthermore, a
217 concurrent paucity of non-pathogenic, commensal *Lactobacillus* species in BV+ patients was
218 also observed. Together, these patterns reinforce one another in the diagnosis of BV by CLS2.0q

219 consistent with *a priori* biological expectations and previous analyses (FIG 1a). In contrast,
220 women diagnosed as BV- show precisely the opposite pattern. Specifically, BV- patients
221 generally present with relatively high abundances of multiple putatively-beneficial lactobacillus
222 spp. and a concurrent lack of pathogenic bacteria, especially those most often found at high
223 levels in BV+ patients (e.g. Megasphaera1, BVAB2; FIG 1a). Together, these patterns constitute
224 the primary signal detected by CLS2.0q in diagnosing BV in clinically symptomatic and
225 asymptomatic women.

226 Interestingly, the non-pathogenic bacterium *L. iners* is relatively abundant in many
227 patients regardless of their diagnosis. This suggests either that the presence and abundance of *L.*
228 *iners* is not relevant to BV diagnosis or, more interestingly, that it plays differing roles in the
229 disease- versus healthy-state communities of bacteria. Indeed, previous studies have suggested
230 that *L. iners* acts as a "sentinel" or "transition" species early in the course of pathogenic infection
231 (37) and may persist at relatively high abundance thereafter. Regardless, this pattern of
232 abundance across patients likely does not contribute significantly to the diagnostic signal.
233 Similarly, *G. vaginalis*, *P. bivia*, and *U urealyticum* are often found in both BV+ and BV-
234 patients. This also suggests they do not contribute strongly to the diagnosis of disease. Rather,
235 this pattern likely reflects the natural dynamics of a community of bacteria at equilibrium with
236 beneficial types, keeping the explosive growth of pathogenic organisms in check.

237 Several, relatively rarely observed, pathogenic bacteria (e.g. Megasphaera 2 and *M.*
238 *mulieris*) were included in testing. Despite this rarity, it is clear that they contribute to the
239 overall patterns discussed above. In point of fact, their strongest contribution to diagnosis is
240 likely their conspicuous absence in BV- patients.

241 Over the course of data analysis, 20 subjects with intermediate Nugent scores were
242 identified. Such patients represent an ambiguous class in current, diagnostic practice with
243 unclear guidelines for treatment. Interestingly, via the mutually reinforcing patterns described
244 above, CLS2.0q appears to provide clear resolution for such patients (Fig. 1b). Specifically,
245 roughly half (n=9) are diagnosed as BV+ by the algorithm and half (n=11) are diagnosed as BV-.
246 Upon further inspection, all asymptomatic women in the validation dataset (N = 23) show
247 some evidence of BV by Nugent analysis. Specifically, 16 score as "High" and the remaining
248 seven as intermediate upon scoring. None score as "Low." Interestingly, four of these women
249 are diagnosed as BV+ by CLS2.0q, all of which are Nugent intermediate but were designated by
250 the microscopist as having "few" clue cells.
251 Assays for candidiasis, trichomoniasis, and five specific *Candida* species were included
252 in the final panel design to enable differential diagnosis for these infections due to their clinically
253 similar presentations (TABLE 2).
254

255 ## Discussion

256 With a measured prevalence between 29% and 50% in women of reproductive age, there
257 are a significant number of women who are negatively impacted by bacterial vaginosis infections
258 each year. Infection is associated with serious potential side effects including increased risk of
259 preterm birth in women affected during pregnancy. However, the complex etiology of the
260 condition — the way multiple distinct combinations of bacteria are associated with a disrupted
261 vaginal microbiome — has only recently begun to be appreciated. Furthermore, patients present
262 with a wide range of clinical symptoms from entirely asymptomatic to showing multiple signs
263 and symptoms. Others are actually infected with other pathogens that cause similar signs and

264 symptoms. The dynamic and highly interactive nature of the underlying cause of disease and
265 complicated clinical presentation makes accurate diagnosis—and differential diagnosis—a
266 challenge.

267 As such, improved strategies for detecting BV infection have been a subject of much
268 discussion. Recent advances in sequencing technologies, the very same that have begun
269 revealing the biology responsible for disease, also show great promise for improving diagnosis.
270 They even open the door to the development of personalized treatment plans based on the actual
271 combination of bacteria at both increased and decreased relative abundance.

272 Here we describe a novel, molecular diagnostic based on quantitative measurement of the
273 relative abundance of multiple pathogenic and commensal organisms of the vaginal microbiome.
274 It does not rely on culture, microscopy, or subjective analysis of patterns of staining. Rather,
275 mutually reinforcing patterns in the observed levels of each of the assayed microorganisms
276 accurately capture the biological complexities of the vaginal microbiome and clearly
277 differentiate BV positive and BV negative patients.

278 The complicated biology underlying BV infections suggests that accurate methods for
279 diagnosis must take into account the balance of multiple organisms simultaneously. Historically,
280 methods such as Amsel's criteria have done this via numerous clinical observations and tests, or
281 by bacterial staining and other laboratory-based techniques in the case of Nugent scoring.
282 However, these methods utilize presence/absence determinations for some diagnostic features
283 and for others provide merely a semi-quantitative categorization of patients. These limitations
284 result in poor resolution by the gold-standard approach. Specifically, Nugent scoring provides a
285 somewhat-ambiguous finding of "Low," "Intermediate," or "High" rather than a more concrete
286 diagnosis of "BV+" or "BV-."

287 The diagnostic method we validate here appears to much more accurately capture the
288 complex biology of BV compared to existing alternatives. It identifies a clear reinforcing pattern
289 of abundance (see Fig. 1a) of multiple pathogenic and commensal *Lactobacilli* species that
290 provides a clear and accurate diagnosis, even in diagnostically ambiguous "Intermediate"
291 patients (Fig. 1b). The method also provides improved sensitivity and specificity over current
292 standard-of-care diagnostic tools. Interestingly, these metrics likely underestimate the number of
293 true positive and true negative individuals, since the algorithm was necessarily compared to a
294 standard with known flaws. It is possible that these flaws produce an upper limit on performance
295 that is unrelated to the clear signal produced by CLS2.0q. The actual sensitivity and specificity
296 of the CLS2.0q diagnostic algorithm may therefore be greater than reported.
297 In addition to providing clarity for the ambiguous Nugent "Intermediate" subjects,
298 CLS2.0q also appears to perform well in asymptomatic patients. This is especially valuable as it
299 is well documented that a majority of BV cases present with no clinical symptoms at all.
300 Interestingly, most of the asymptomatic patients involved in the validation were diagnosable by
301 CLS2.0q as BV+ (FIG 1c). Again, via clearly interpretable biological pattern recognition,
302 CLS2.0q appears to aid in the accurate diagnosis of such patients as well, though additional
303 studies focused on this subgroup would be necessary to confirm this observation.
304 In addition to the improved capture of biologically-consistent, patient-specific patterns of
305 dysregulation in the vaginal microbiome, CLS2.0q also incorporates several technical
306 improvements over existing methods. For example, together with the detection technology, it is
307 capable of identifying both rare and difficult-to-culture organisms that are currently missed by
308 standard diagnosis methods. Additionally, OpenArray technology provides several key
309 advantages over previous approaches. These include increased flexibility of microbial target

310 choices, improved accuracy of microbial identification, and high-throughput capability. They
311 also allow for the refinement of platform content as future research continues to reveal the
312 underlying dynamics of infection.

313 This successful validation of the CLS2.0q algorithm was nonetheless limited in several
314 ways. The first is the somewhat small overall sample size in the validation set of 172 women.
315 However, though the validation dataset may seem limited, the original training set included 410
316 symptomatic and asymptomatic women, and all findings during the original algorithm
317 development were very similar to those presented here (data not shown). The second limitation
318 is this study's somewhat biased sample of patients. Whereas the majority of subjects in this
319 dataset self-identify as Latina, the ideal validation would have included representative samples of
320 all ethnicities in order to fully control for known and potential differences amongst groups. An
321 ideal study would also have fully controlled for geography and socio-economic status. However,
322 the availability of subjects and other logistical challenges limited collection to a single
323 location/health system whose population may not be fully representative at the national level.

324 Another limitation is inherent to the biology of the condition itself, namely the dynamics
325 of the vaginal microbiome in real-time that inevitably introduces variation due to the status of the
326 dysregulation at the time of sampling. While formally unaccounted for in our analyses, the
327 robustness of the algorithm displayed in development and in this validation strongly suggest that
328 we have sampled over a sufficient number of dysregulated states to reflect the majority of the
329 relevant biological variation detectable by the method.

330 Ultimately, this validation established CLS2.0q as a quantitative, sensitive, specific,
331 accurate, and robust algorithm for the diagnosis of bacterial vaginosis. It is demonstrably
332 superior to gold-standard approaches at identifying the underlying biological dynamics of

333 dysregulation in the vaginal microbiome that underlie the condition. It enables efficient
334 differential diagnosis of multiple other conditions that present in a clinically similar manner.
335 Finally, future work may allow the algorithm to be extended to the data-driven recommendation
336 of personalized treatment regimes, including those that may help circumvent and/or avoid
337 antibiotic resistance. Ongoing research is expected to identify additional organisms of interest
338 for bacterial vaginosis, and CLS2.0q is well positioned to rapidly include this information in
339 future refinements of the algorithm.

340 It is evident that quantitative methods of diagnosis provide key improvements over
341 current approaches to the diagnosis and treatment of complex infections. It is hoped that the
342 technology employed here will enable point-of-care compatible turnaround times and broaden
343 the impact of these techniques in personalized medicine. In addition to flexibility and efficiency,
344 there is also the potential for these approaches to identify drug-resistant organisms, and thus to
345 suggest specific and personalized treatment regimens for individual patients. Furthermore, these
346 techniques can be used to test for and diagnose other conditions, such as respiratory infections,
347 gastrointestinal dysregulation, and wound healing concerns.

348

349 Acknowledgments

350 The authors thank ThermoFisher Scientific for their support in obtaining the data used to train
351 and develop the CLS2.0q algorithm.

352 # References

353 1. Sha BE, Chen HY, Wang QJ, Zariffard MR, Cohen MH, Spear GT. 2005. Utility of Amsel
354 criteria, Nugent score, and quantitative PCR for *Gardnerella vaginalis, Mycoplasma hominis,*
355 and *Lactobacillus* spp. for diagnosis of bacterial vaginosis in human immunodeficiency virus-
356 infected women. J Clin Microbiol 43:4607-12.

357 2. Allsworth JE, Peipert JF. 2007. Prevalence of bacterial vaginosis: 2001-2004 National Health
358 and Nutrition Examination Survey data. Obstet Gynecol 109:114-120.

359 3. Yen S, Shafer MA, Moncada J, Campbell CJ, Flinn CD, Boyer CB 2003. Bacterial vaginosis
360 in sexually experienced and non-sexually experienced young women entering the military.
361 Obstet Gynecol 102:927-33.

362 4. Fethers KA, Fairley CK, Morton A, Hocking JS, Hopkins C, Kennedy LJ, Fehler G, Bradshaw
363 CS. 2009. Early sexual experiences and risk factors for bacterial vaginosis. J Infect Dis
364 200:1662-1670.

365 5. Esber A, Miguel RDV, Cherpes TL, Klebanoff MA, Gallo MF, Turner AN. 2015. Risk of
366 bacterial vaginosis among women with herpes simplex virus type 2 infection: a systematic
367 review and meta-analysis. J Infect Dis 212:8-17.

368 6. Flynn CA, Helwig AL, Meurer LN. 1999. Bacterial vaginosis in pregnancy and the risk of
369 prematurity: a meta-analysis. J Fam Pract 48:885-892.

370 7. Myer L, Denny L, Telerant R. 2005. Bacterial vaginosis and susceptibility to HIV infection in
371 South African women: a nested case-control study. J Infect Dis 192:1372-1380.

372 8. Bilardi J, Walker S, McNair R, Mooney-Somers J, Temple-Smith M, Bellhouse C, Fairley C,
373 Chen M, Bradshaw C. 2016. Women's management of recurrent bacterial vaginosis and
374 experiences of clinical care: a qualitative study. PLoS ONE 11:e0151794.

375 9. Cartwright CP, Lembke BD, Ramachandran K, Body BA, Nye MB, Rivers CA, Schwebke JR.
376 2012. Development and validation of a semiquantitative, multitarget PCR assay for diagnosis of
377 bacterial vaginosis. J Clin Microbiol 50:2321-29.

378 10. Oakley BB, Fiedler TL, Marrazzo JM, Fredricks DN. 2008. Diversity of human vaginal
379 bacterial communities and associations with clinically defined bacterial vaginosis. J Appl
380 Environ Micro 74:4898-4909.

381 11. Koumans EH, Sternberg M, Bruce C, McQuillan G, Kendrick J, Sutton M, Markowitz LE.
382 2007. The prevalence of bacterial vaginosis in the United States, 2001-2004; associations with
383 symptoms, sexual behaviors, and reproductive health. Sex Transm Dis 34:864-869.

384 12. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, Karlebach S, Gorle R,
385 Russell J, Tacket CO, Brotman RM, Davis CC, Ault K, Peralta L, Forney LJ. 2010. Proc Natl
386 Acad Sci USA 108 (suppl 1):4680-87.

387 13. Nugent RP, Krohn MA, Hillier SL. 1991. Reliability of diagnosing bacterial vaginosis is
388 improved by a standardized method of Gram stain interpretation. J Clin Microbiol 29:297-301.

389 14. Menard JP, Fenollar F, Henry M, Bretelle F, Raoult D. 2008. Molecular quantification of
390 *Gardnerella vaginalis* and *Atopobium vaginae* loads to predict bacterial vaginosis. Clin Infect
391 Dis 47:33-43.

392 15. Fredricks DN, Fiedler TL, Thomas KK, Oakley BB, Marrazzo JM. 2007. Targeted PCR for
393 detection of vaginal bacteria associated with bacterial vaginosis. J Clin Microbiol 45:3270-76.

394 16. Amsel R, Totten PA, Spiegel CA, Chen KC, Eschenbach D, Holmes KK. 1983. Nonspecific
395 vaginitis: diagnostic criteria and microbial and epidemiological associations. Am J Med 74:14-
396 22.
397 17. Shipitsyna E, Roos A, Datcu R, Hallen A, Fredlund H, Jensen JS, Engstrand L, Unemo M.
398 2013. Composition of the vaginal microbiota in women of reproductive age – sensitive and
399 specific molecular diagnosis of bacterial vaginosis is possible? Plos ONE 8:e60670.
400 18. Koumans EH, Markowitz LE, Hogan V. 2002. Indications for therapy and treatment
401 recommendations for bacterial vaginosis in nonpregnant and pregnant women: a synthesis of
402 data. Clin Infect Dis 35: S152-S172.
403 19. Briselden AM, Hillier SL. 1994. Evaluation of Affirm VP Microbial Identification Test for
404 *Gardnerella vaginalis* and *Trichomonas vaginalis*. J Clin Microbiol 32:148-52.
405 20. Bradshaw CS, Morton AN, Garland SM, Horvath LB, Kuzevska I, and Fairley CK. 2005.
406 Evaluation of a point-of-care test, BVBlue, and clinical and laboratory criteria for diagnosis of
407 bacterial vaginosis. J Clin Microbiol 43:1304-08.
408 21. Gaydos CA, Beqaj S, Schwebke JR, Lebed J, Smith B, Davis TE, Fife KH, Nyirjesy P,
409 Spurrell T, Furgerson D, Coleman J, Paradis S, Cooper CK. 2017. Clinical validation of a test for
410 the diagnosis of vaginitis. Obstet Gynecol 130:181-189.
411 22. Mills BB. 2017. Vaginitis: beyond the basics. Clin Obstet Gynecol N Am, 44:159–177.
412 23. Ling Z, Kong J, Liu F, Zhu H, Chen X, Wang Y, Li L, Nelson K, Xia Y, Xiang C. 2010.
413 Molecular analysis of the diversity of vaginal microbiota associated with bacterial vaginosis.
414 BMC Genomics 11:488.

415 24. Srinivasan S, Morgan MT, Liu C, Matsen FA, Hoffman NG, Fiedler TL, Agnew KJ,
416 Marrazzo JM, Fredricks DN. 2013. More than meets the eye: associations of vaginal bacteria
417 with Gram stain morphotypes using molecular phylogenetic analysis. PLoS ONE 8:e78633.
418 25. Tamrakar R, Yamada T, Furuta I, Cho K, Morikawa M, Yamada H, Sakuragi N, Minakami
419 H. 2007. Association between *Lactobacillus* species and bacterial vaginosis-related bacteria, and
420 bacterial vaginosis scores in pregnant women. BMC Infect Dis 7:128.
421 26. Bradshaw CS, Tabrizi SN, Fairley CK, Morton AN, Rudland E, Garland SM. 2006. The
422 association of *Atopobium vaginae* and *Gardnerella vaginalis* with bacterial vaginosis and
423 recurrence after oral metronidazole therapy. J Infect Dis 194:828-836.
424 27. Zozaya-Hinchliffe M, Lillis R, Martin DH, Ferris MJ. 2010. Quantitative PCR assessments
425 of bacterial species in women with and without bacterial vaginosis. J Clin Microbiol 48:1812-19.
426 28. Srinivasan S, Fredricks D. 2008. The human vaginal bacterial biota and bacterial vaginosis.
427 Interdiscip Perspect Infect Dis 2008:e750479.
428 29. Verstraelen H, Verhelst R. 2009. Bacterial vaginosis: an update on diagnosis and treatment.
429 Expert Rev Anti Infect Ther 7:1109-24.
430 30. Lamont RF, Sobel JD, Akins RA, Hassan SS, Chaiworapongsa T, Kusanovic JP. 2011. The
431 vaginal microbiome: new information about genital tract flora using molecular based techniques.
432 BJOG 118:533-49.
433 31. van de Wijgert JHHM, Borgdorff H, Verhelst R, Crucitti T, Francis S, Verstraelen H, Jespers
434 V. 2014. The vaginal microbiota: what have we learned after a decade of molecular
435 characterization? PLoS ONE 9: e105998.
436 32. Martin DH, Marrazzo JM. 2016. The vaginal microbiome: current understanding and future
437 directions. J Infect Dis 214:S36-S41.

438 33. Datcu R, Gesink D, Mulvad G, Montgomery-Anderson R, Rink E, Koch A, Ahrens P, Jensen
439 JS. 2013. Vaginal microbiome in women from Greenland assessed by microscopy and
440 quantitative PCR. BMC Infect Dis 13:480.
441 34. Hilbert DW, Smith WL, Paulish-Miller TE, Chadwick SG, Toner G, Mordechai E, Adelson,
442 ME, Sobel JD, Gygax SE. 2016. Utilization of molecular methods to identify prognostic markers
443 for recurrent bacterial vaginosis. Diagn Microbiol Infect Dis 86:231-42.
444 35. Plummer EL, Garland SM, Bradshaw CS, Law MG, Vodstrcil LA, Hocking JS, Fairley CK,
445 Tabrizi SN. 2017. Molecular diagnosis of bacterial vaginosis: Does adjustment for total bacterial
446 load or human cellular content improve diagnostic performance? J Microbiol Methods 133:66-
447 68.
448 36. Rivers CA, Lee JY, Sharples N, Ledeboer NA, Schwebke JR. 2014. ESwab™ as an optional
449 collection device for use with the Affirm VPIII Microbial Test System. J Clin Microbiol
450 52:1698-1700.
451 37. Jakobsson T, Forsum U. 2007. *Lactobacillus iners*: a marker of changes in the vaginal flora?
452 J Clin.Microbiol.45:3145.

453

454 Tables & Figures

455

Pathogenic	Commensal lactobacilli
Atopobium vaginae	*Lactobacillus crispatus*
*BVAB2**	*Lactobacillus gasseri*
Gardnerella vaginalis	*Lactobacillus iners*
Megasphaera 1	*Lactobacillus jensenii*
Megasphaera 2	
Mobiluncus curtisii	
Mobiluncus mulieris	
Mycoplasma hominis	
Prevotella bivia	
Ureaplasma urealyticum	

*Bacterial Vaginosis–Associated Bacterium type 2

456 **TABLE 1: Microorganisms included in the CLS2.0q molecular assay for Bacterial**
457 **Vaginosis diagnosis.** After extensive literature review, these ten likely-pathogenic and four
458 commensal bacteria were chosen as targets associated with bacterial vaginosis. The relative
459 amounts of each of these bacteria present in a patient sample inform the CLS2.0q diagnostic.

460

Candidiasis	
Candida albicans	*Candida krusei*
Candida parapsilosis	*Candida tropicalis*
Candida glabrata	

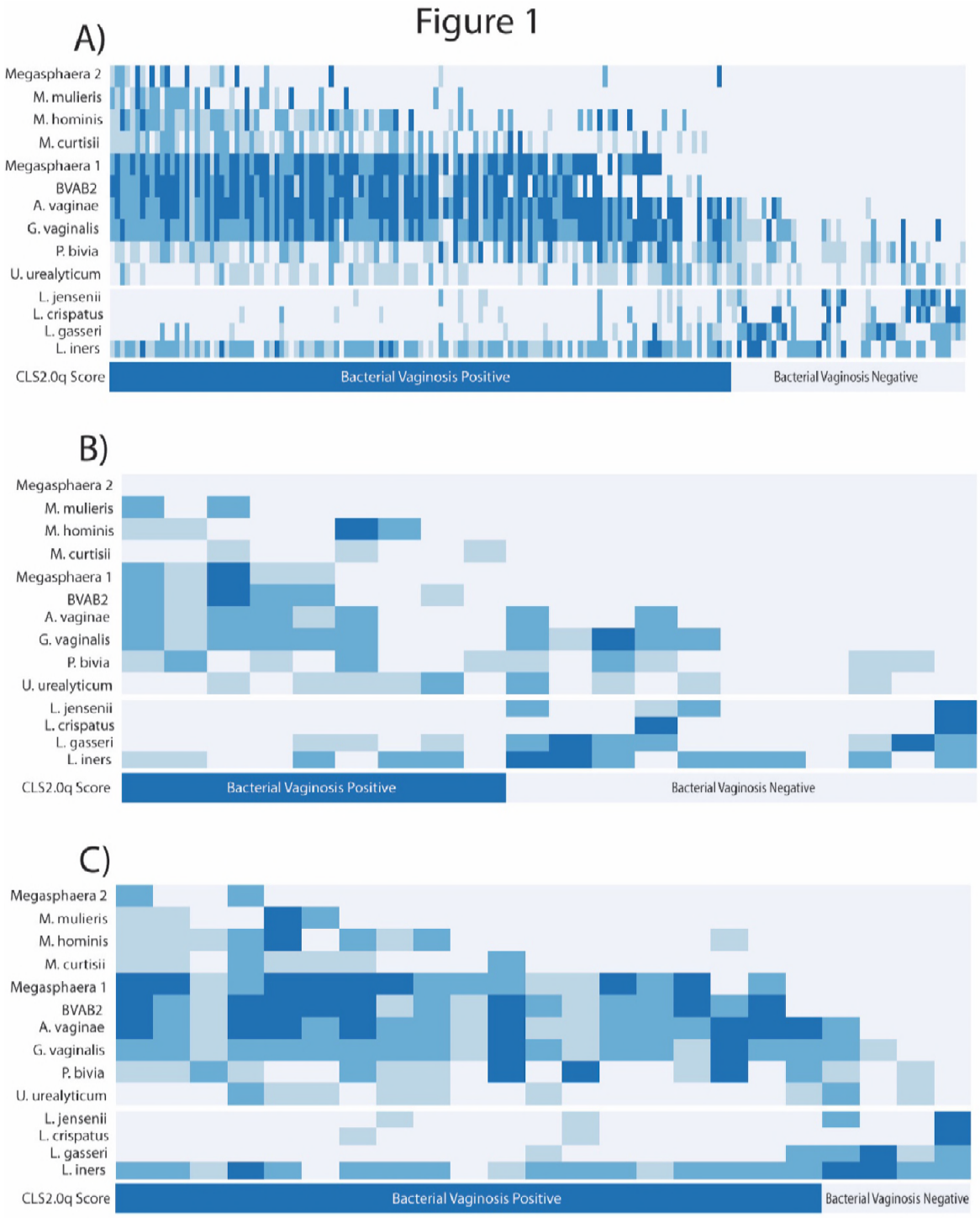
Sexually Transmitted infections
Trichomonas vaginalis

461

462 **TABLE 2: Additional organisms included in the CLS2.0q molecular assay for differential**
463 **diagnosis.** After extensive literature review these organisms were chosen as targets associated
464 with infections which present with clinically similar signs and symptoms to BV. The presence of
465 any of these informs the process of differential diagnosis.

466

467 FIG 1A:



468

469

470

471 **FIG1: Relative bacterial abundance in the validation data set (N = 172), and its role in**
472 **CLS2.0q diagnosis of bacterial vaginosis in vaginal microbiome.** (A) Relative bacterial
473 abundance (rows) of pathogenic (top section), commensal Lactobacilli (middle section) in
474 women (columns) diagnosed as BV+ and BV- (bottom section) by CLS2.0q in the validation
475 dataset. No single organism perfectly differentiates BV+ and BV- patients. Rather, BV+
476 subjects show relatively high levels of multiple pathogenic bacteria especially Megasphaera 1,
477 BVAB2, *A. vaginae*, and *G. vaginalis* (top-left) and relatively low levels of commensal
478 Lactobacillus species (middle-left). In contrast, BV- subjects show relatively high abundances of
479 multiple Lactobacillus spp. (middle-right) and a concurrent lack of pathogenic bacteria (top-
480 right), especially Megasphaera1 and BVAB2. Together, these mirror-image patterns constitute
481 the primary signal detected by CLS2.0q in diagnosing BV in clinically symptomatic and
482 asymptomatic women.
483 (B) Relative bacterial abundance and CLS2.0q diagnosis of BV for samples with a Nugent
484 categorization of "Intermediate" (N = 20). A strikingly similar pattern as described above is
485 associated with BV diagnosis by CLS2.0q.
486 (C) Relative bacterial abundance and CLS2.0q diagnosis of BV for asymptomatic individuals (N
487 = 23). A similar and consistent pattern as described above is also associated with BV diagnosis
488 by CLS2.0q.