Title: Diagnostic-avoiding *Chlamydia trachomatis* variants detected in cervical and endometrial specimens from women during 16S microbiome profiling.

Authors: Sangmi Jeong[1], Tammy Tollison[1], Hayden Brochu[1], Hsuan Chou[1], Tammy Yu[2], Priyanka Baghaie[2], Kacy S. Yount[2], Toni Darville[2], Harold C Wiesenfeld[3], Sharon L Hillier[3], Xinxia Peng[1] and Catherine M. O'Connell[2]

[1]North Carolina State University

[2]University of North Carolina at Chapel Hill

[3]University of Pittsburgh School of Medicine and the Magee-Womens Research Institute

Correspondence: Catherine O'Connell, PhD, University of North Carolina at Chapel Hill, 8340B MBRB, CB# 7509, 111 Mason Farm Road, Chapel Hill, NC 27599-7509, phone (919) 962-3422, fax:(919) 966-7277, catherine.oconnell@unc.edu

Alternate corresponding author: Xinxia Peng, PhD, North Carolina State University, Raleigh, NC 27695, phone (919) 515-4481, fax: (919)-513-6464, xpeng5@ncsu.edu

Keywords: NAAT, chlamydia, gonorrhea, microbiome, 16S rRNA, false negative, 23S

Abstract:

Background: Performance of a 16S rRNA analysis of the cervicovaginal microbiome of 220 participants recruited into the T Cell Response against Chlamydia (TRAC) cohort between February 2011 and August 2014 in Allegheny County, Pennsylvania USA detected DNA encoding chlamydial 16S rRNA in samples from seven participants whose tests were negative for *Chlamydia trachomatis* (CT) and DNA encoding gonococcal 16S rRNA from five participants whose tests were negative for *Neisseria gonorrhoeae* (NG) infection with the Aptima Combo2 assay (Hologic).

Methods: We used targeted PCR amplification followed by sequencing to characterize the chlamydial 23S rRNA locus and qPCR to detect gonococcal DNA in residual diagnostic swab eluates or DNA used to generate 16S rRNA libraries.

Results: Discrepant specimens that contained chlamydial DNA carried a diagnostic-avoidant, G1526A variant in the 23S rRNA locus identical to variants previously detected in Finland, Denmark, and the UK. PCR validation of gonococcal DNA was confirmed for all participants whose tests were negative, with stochastic effects consistent with infection levels close to the limit of detection by the diagnostic assay.

Conclusions: These data indicate that this probe-avoidant CT mutant, and possibly others, were circulating in the northeastern US prior to their detection and characterization in 2019. Although infrequent, documentation of false negative results for CT indicates a need for clinicians to consider performance of a second test that uses alternate PCR targets if patients have persistent symptoms or have known contact to an infected sex partner and their initial NAAT is negative.

41 **Background**:

42 Performing cross sectional or longitudinal research studies investigating natural immunity
43 or immunopathogenesis of sexually transmitted infections (STI) in humans can be a complex,
44 time consuming and expensive undertaking. Efforts to mine all data associated with infection or
45 disease cohorts are often extensive and protracted. Such is the case for the TRAC (T cell
46 Response Against Chlamydia) cohort, a group of 246 cisgender women at elevated risk for
47 chlamydial infection who enrolled between February 2011 and August 2014 while presenting to
48 the Allegheny County, PA Health Department's Sexually Transmitted Diseases Clinic for STI
49 testing and/or management. The TRAC study sought to profile cell mediated responses
50 associated with protection against ascending *C. trachomatis* (CT) infection (from the cervix to
51 endometrium) or against CT reinfection [1, 2]. Since completion of enrollment, data obtained from
52 participant samples and questionnaire responses have been used to profile antibody [1, 3-5], T
53 cell antigen-specific responses [2], and to determine risk factors for ascending infection. Using
54 modern molecular techniques, we quantified pathogen burden [1, 2], local cytokine secretion [6]
55 and used genetic profiling techniques to search for biomarkers predictive of disease or reinfection
56 risk [7-9]. Secondary investigations have also assessed the impact of *Mycoplasma genitalium*
57 infection in this population [10]. Fundamental to the success of these analyses has been accurate
58 classification of study participants according to infection status at enrollment and during
59 subsequent follow up, so enrollee infection status was assigned according to the outcome of an
60 FDA-cleared STI diagnostic NAAT test performed at the study visit, the results of which were used
61 to inform clinical care. Infrequently, discrepant results were observed between research PCR
62 assays detecting chlamydial loci and results obtained using the Aptima Combo 2® assay (AC2),
63 manufactured by Hologic, Inc., a commercially available nucleic acid amplification test (NAAT)
64 for CT and *Neisseria gonorrhoeae* (NG), targeting 23S ribosomal RNA (rRNA) and 16S rRNA in
65 these bacteria, respectively. However, we could not exclude the possibility of cross-contamination

66 from the environment or chlamydia-infected specimens, particularly with repurposed, reprocessed
67 convenience samples, e.g., leftover transport medium.

68 The availability of pristine “future use” vaginal swabs collected from the same cohort and
69 stored since enrollment provided an opportunity to reexamine these discrepant results while
70 determining cervicovaginal microbiome profiles that associated with ascending chlamydial
71 infection [11]. We performed a 16S rRNA gene genomic DNA based microbiome analysis on the
72 entire cohort and detected chlamydial DNA in samples from seven participants whose tests were
73 negative for CT and gonococcal DNA from five participants whose tests were negative for NG
74 with the AC2 assay. PCR amplification and subsequent sequencing using primers directed
75 against the 23S rRNA locus revealed that some of the specimens contained chlamydial DNA
76 carrying a diagnostic-avoidant single nucleotide polymorphism (SNP) that likely contributed to
77 misclassification. These data indicate that this mutant, and possibly others, were circulating in the
78 northeastern US prior to their detection and characterization in the United Kingdom and Denmark
79 in 2019 and subsequently [12-14]. Discrepant NG detection reflected very low gonococcal loads
80 that challenged the limit of resolution for the NAAT. Although infrequent, documentation of false
81 negative results for CT or NG may indicate a need for clinicians to consider performance of a
82 second test that uses alternate or additional PCR probes.

83

84 **Methods:**

85 *T Cell Response against Chlamydia (TRAC) Cohort (246 participants)*. This cohort was composed
86 of young (median age, 21 years; range, 18-35 years) cis-gender women at high risk for the
87 acquisition of CT [1]. Criteria indicating high-risk status included > 3 sexual partners in the
88 previous 6 months, ≤ 14 years of age at sexual debut, history of pelvic inflammatory disease
89 (PID), the presence of mucopurulent cervicitis on exam, or sexual contact with a partner known

90 to be infected with CT or NG or non-gonococcal non-chlamydial urethritis. Women with a current
91 diagnosis of PID were excluded. Additional exclusion criteria were pregnancy, uterine procedure
92 or miscarriage in the preceding 60 days, menopause, hysterectomy, antibiotic therapy in the
93 preceding 14 days, and allergy to study medications. TRAC participants were enrolled into a
94 longitudinal study designed to investigate T cell responses important for protection from incident
95 chlamydial infection over 12 months of follow-up between February 2011 and August 2014 and
96 were recruited from the Allegheny County Health Department's Sexually Transmitted Diseases
97 Clinic, UPMC Magee-Womens Hospital (MWH) Ambulatory Care Clinic, and the Reproductive
98 Infectious Disease Research Unit at MWH in Pittsburgh PA. Participants provided informed
99 consent at the time of enrollment and completed questionnaires regarding obstetric/gynecologic
100 history, behavioral practices, sex exposure, contraceptive methods, and symptoms. Clinical,
101 histological, and microbiological testing was performed, and blood and endometrial samples were
102 obtained after enrollment, after which all participants received single-dose agents for gonorrhea
103 (ceftriaxone, 250 mg intramuscularly) and chlamydia (azithromycin, 1 g orally). Participants in this
104 cohort were assessed for cervical and endometrial infection using AC2 (Hologic, Marlborough,
105 MA) NAAT with overall infection rates of 67% and 8.5% for CT and NG, respectively. *M. genitalium*
106 (MG) infection was determined using the Aptima MG NAAT (Hologic).

107

108 *16S library construction and characterization.* Cervicovaginal swab samples available from 246
109 women enrolled in TRAC were processed for 16S rRNA gene microbiome profiling [11]. In brief,
110 cervicovaginal microbial communities present in participant specimens were profiled by
111 sequencing the V4 region of 16S rRNA gene using Illumina sequencing. The 16S rRNA amplicon
112 sequencing data was processed at the level of inferred amplicon sequence variants (ASVs) using
113 USEARCH v11 [15]. The raw sequencing data supporting the conclusions of this article is publicly

114 available at NCBI's Short Read Archive (SRA) under BioProject accession number
115 PRJNA1136868.
116

117 *PCR template purification from residual diagnostic samples.* DNA was extracted as previously
118 published using a Quick-DNA Universal Kit (Zymo Research, Irvine, California, USA) from 100 µL
119 to 300 µL of reserved cervical swab eluates. Chlamydial and/or mycoplasmal loads were
120 measured by quantitative PCR [1] targeting the 16SrRNA gene [16] or the highly conserved
121 MG190 locus [10], respectively. Primers directed against the *parA* pseudogene [17], 16S rRNA
122 and 23S rRNA loci [18] were similarly used to validate and/or quantify NG abundance.
123

124 *PCR amplification and sequencing of 23S rRNA alleles*. Amplification conditions for detection of
125 SNPs in *C. trachomatis* 23S were as described [19]. The melting temperature was raised to 98°C
126 for optimal activation of the high fidelity, proof-reading enzyme Phusion (ThermoFisher). Samples
127 were processed independently in a dedicated PCR hood to minimize risk of cross-contamination
128 during analysis. Two rounds of amplification were required to recover sufficient product which was
129 cleaned and concentrated using a DNA Clean and Concentrate kit (Zymo) before sequencing in
130 forward and reverse directions using V3 region primers [19]. Two patient samples that had tested
131 positively using the AC2 test were also analyzed as controls. Sequences generated via Sanger
132 sequencing were aligned against the published sequence using Clone Manager Professional
133 Version 11 (Sci Ed Software LLC, Westminster, CO). Residual cervical templates were used from
134 profiled participants except for 3-209, for whom only endometrium-extracted DNA remained
135 available for analysis.

136 *Statistical analysis*. The nonparametric Spearman rank correlation coefficient was calculated to
137 assess the overall correlation between CT 16S rRNA gene DNA copy numbers measured by
138 qPCR and the raw read counts of CT 16S rRNA gene from microbiome sequencing analysis using

139 R function ‘cor’ [20]. The LOESS (Locally Weighted Scatter-plot Smoother) curve was estimated
140 and plotted using a ‘geom_smooth’ function of the R ‘ggplot2’ package with default settings [21].
141 Abundance of NG-specific ASVs present in libraries from individuals testing positively with AC2
142 and those who tested negative were compared using a Mann-Whitney test (two tailed).

143

144 **Results:**

145 *Detection of chlamydia-derived amplicon sequence variant (ASV) in 16S libraries derived from*
146 *vaginal swabs obtained from CT negative TRAC participants.* Total DNA was extracted from
147 “future use” vaginal swabs from a total of 246 TRAC participants for inclusion in a 16S rRNA gene-
148 based sequencing analysis of the contribution of the cervicovaginal microbiome to ascending
149 chlamydial infection. Three samples were excluded based on DNA low integrity values after
150 extraction. We also excluded twenty-three samples due to the equivocal results of chlamydial
151 infection at the cervix and/or endometrium from clinical diagnostic testing. The remaining 220
152 samples were processed using the ASV-level analysis, resulting in 1,990 ASVs detected in total
153 from samples sequenced to the average depth of ~151,000 reads. Taxonomy annotations of
154 ASVs were based on sequence comparison against the 16S rRNA gene reference database RDP
155 version 19 [22] using the SINTAX algorithm [23]. We assessed the concordance of STI detection
156 by examining raw read counts of sexually transmitted pathogen-specific 16S rRNA genes vs
157 infections previously assessed by NAAT diagnostic tests (CT, NG, MG) (**Fig. 1a**) and quantified
158 via qPCR (CT [1], **Fig. 1b**). About 58% (83 of 143) samples from participants whose tests were
159 positive for CT at enrollment had non-zero 16S rRNA gene read counts (**Fig. 1a**). Similarly, 85%
160 (17 of 20) samples from participants whose tests were postive for NG, and 15.8% (6 of 38)
161 samples from MG positive enrollees had non-zero 16S rRNA gene read counts. Overall, there
162 was modest correlation (Spearman = 0.575) between the two CT quantification methods (**Fig.**
163 **1b**).

164 Based on clinical diagnostic results obtained at enrollment, the TRAC participants were
165 classified into three groups: 1) Uninfected; 2) Endo- participants with CT infection detected only
166 in their cervix; 3) Endo+ participants who tested positively for CT in cervix and endometrium.
167 There were no TRAC enrollees whose tests were positive only in the endometrial sample.
168 Comparing ASV abundances between CT-positive (Endo- and Endo+) and CT-negative (control)
169 women, we were surprised to detect 16S ASV within libraries derived from samples from women
170 who had been assigned "uninfected" (N=7) based on the results of NAAT testing (**Fig. 1a**). These
171 results correlated with estimated chlamydial abundance derived from a previously performed
172 quantitative 16S PCR assay (**Fig. 1b** and **Table 1**), suggesting that attributing the original results
173 to external or cross-contamination was incorrect. A review of these participants' clinical data
174 revealed that they were mostly asymptomatic (5/7), one had reported experiencing dysuria,
175 consistent with >10 leucocytes detected per high power field in her urine sample with irregular
176 vaginal bleeding and another reported dyspareunia and abnormal vaginal discharge. Clinical
177 findings upon pelvic examination were unremarkable for all. Three of the seven participants within
178 this group reported contact with a CT infected partner in the three months prior to their enrollment
179 but had yet to be treated.

180

Table 1. TRAC participants with discrepant clinical and molecular CT detection

Sample ID	Clinical testing (CT, MG, NG)	16S rRNA read counts in ASV analysis	16S qPCR*	
			Cx	Endo
3-101	CT-/MG-/NG-	1	78	919
3-144		16	420	87
3-145		146	128800	35
3-148		111	3839	457
3-201		55	1495	721
3-226		31	1640	105
3-209	CT-/MG+/NG-	4	13000	62680

*Chlamydial 16S rDNA copies/swab [1].

181

182 In the context of the TRAC cohort, reasons for false negative CT NAAT test results, other
183 than testing failures detected at or during performance e.g., compromised reagents or automated
184 run failure (batch specific), include inadequate sampling at the anatomical site or an unanticipated
185 flaw in the diagnostic test. All NAAT tests were run within temporal parameters appropriate for
186 clinical STI testing, so batch-specific effects did not serve as an explanation for these
187 discrepancies because the participants were not recruited contemporaneously. Testing of an
188 endometrial biopsy might result in a false negative because the entire endometrial lining is not
189 sampled, the amounts of tissue obtained may vary [7], and the region of specimen dedicated to
190 NAAT testing is small [10]. However, cervical testing, particularly when performed by highly
191 trained practitioners, is robust [24, 25]. Several publications from Northern Europe have reported
192 instances of false negative testing specifically associated with the AC2 test [12, 13, 26, 27] arising
193 from probe-avoidant mutations at the chlamydial 23S rRNA locus [19]. Speculating that this might
194 explain our discrepant results, we attempted to amplify the relevant V3 region from cervically or
195 endometrially derived templates that had been extracted from residual diagnostic swab eluates.
196 In all instances we were successful using the protocol and primers used by Hokynar et al. [19],
197 further confirming the presence of chlamydial DNA in the clinical samples. We detected a G-A
198 substitution at 1523 in the samples of all seven test-discrepant participants (**Fig. 2**) when these
199 products were sequenced. NAAT positively-testing samples (N=2) from the cohort were wild type
200 at this position (data not shown). These data indicated that at least one probe-avoidant variant of
201 CT was circulating in the region between May 2012 and February 2014, the time window between
202 enrollments of our earliest and latest discrepant TRAC participants.

203 16S rRNA gene microbiome analysis also revealed a total of five participants with
204 discrepant NG AC2 testing enrolled from April 2011 to July 2013 (**Fig. 1A, Table 2**), including one
205 participant, 3-209, who also tested discrepantly for CT (**Table 1**). One participant reported mild

206 urinary symptoms with abnormal vaginal discharge and irregular bleeding and one other
207 participant reported bleeding with intercourse. Clinical findings on pelvic examination were
208 noteworthy for bleeding with cervical contact for four of five participants but were otherwise
209 unremarkable. Only one of these participants reported contact with an infected partner within the
210 preceding three months but had not received treatment.

211 The genus *Neisseria* includes a group of commensal species, *N. subflava*, *N. mucosa*, *N.*
212 *elongata*, in addition to the important human pathogens, *N. gonorrhoeae* and *N. meningitidis* with
213 the potential for genetic exchange between these naturally competent microorganisms [28] so we
214 initially speculated that the ASV detected lacked discriminatory capacity. However, alignment of
215 the 16S V4 sequence from *N. gonorrhoeae* with the 35 *Neisseria* species included in the RDP
216 v18 database, revealed that ASV ZOtu95, detected in the discrepant samples, was highly specific,
217 aligning with 100% identity between bases 535-787 of the 1544 nucleotide V4 region. This
218 suggested that the 16S ASV detected in libraries generated from the vaginal swabs reflected true
219 gonococcal infection. Next, we compared the number of ZOtu95 reads from discordant individuals
220 with those detected in libraries generated from TRAC participants with concordant NG detection
221 (Table 2) as a surrogate for pathogen load. The number of reads counted in NG+ samples ranged
222 across four orders of magnitude but participant swab samples yielding libraries with ≤6 counts still
223 tested positively via AC2 and no difference was detected between the groups using the non-
224 parametric Mann-Whitney test (P=0.4116). Low levels of infection may have contributed to
225 differential sampling of bacterial populations present at the cervix (diagnostic) and in the vagina
226 (16S analysis). Quantifying NG by qPCR in residual template was difficult because of the minimal
227 amount of cervical DNA remaining and the potential for high genetic variability between strains.
228 Using primers that targeted multicopy (16S, 23S rRNA) and single copy (ParA pseudogene) loci,
229 with genomic DNA extracted from *N. gonorrhoeae* strain FA19 as a positive control [29], we
230 performed qPCR on diagnostic swab eluates or the 16S template that remained from discrepant

231 NG+ participants (Table 2). Evaluating the performance of the primer pairs using material
232 obtained from participants testing positively for NG infection, we observed that only 8 of the 17
233 samples consistently yielded PCR product for all 3 loci, while 6, were positive for 1-2. We were
234 unsuccessful in amplifying material from the remaining 3 samples. Overall, successful
235 amplification of the targets did not reflect abundance of 16S read counts in participants' 16S rRNA
236 libraries. For samples derived from participants who had tested negatively clinically but with
237 ZOtu95 reads detected via 16S rRNA microbiome profiling, all tested positively for at least one
238 multicopy locus.

239 With no source for cross-contamination available in the lab performing the 16S rRNA
240 analysis, we also looked retrospectively at participant enrollment dates to exclude the remote
241 possibility of cross- or environmental contamination [30] occurring during sample collection or
242 NAAT testing. In only one instance did we observe enrollment of a discrepant NG- (3-121) and a
243 confirmed clinically NG+ participant within the same 48-hour window (3-120). However, the
244 presence of additional, neisserial species ASVs in the library generated from 3-120 only (Table
245 2), is inconsistent with contamination of the discrepant NG- sample by material derived from this
246 diagnostic swab.

247

248 **Discussion:**

249 This study revealed that non-congruent CT infection outcomes that we detected between
250 clinical diagnostics and 16S analysis for cervicovaginal samples collected from TRAC enrollees
251 are best explained as resulting from false negative NAAT tests. The point mutation in 23S that we
252 detected in discordant TRAC participants has been associated with reduced probe binding with
253 the original AC2 assay. The G1523A variant was detected in the UK after clinical samples were
254 screened for probe-avoiding mutants that might be in circulation [12] and in Denmark [13]. This is

255 the first report of its detection in the United States although Katz et al., reported other 23S variants
256 that mapped to A1518G or G1526A escaping detection in specimens collected in a five-week
257 window, spanning late 2019 to early 2020 (2 of 401) [31]. Our retrospective analysis suggests
258 that diagnostic-avoidant CT was already circulating in western Pennsylvania, USA by 2012 or
259 possibly earlier. The first participants of the TRAC cohort enrolled in February 2011, but we did
260 not identify diagnostically discordant enrollment samples collected prior to May 2012. Provided
261 that this mutation does not impair chlamydial fitness, it is likely that it continues to circulate. The
262 first 23S SNP variant escaping detection by the AC2 assay, FI-nvCT (C1515T), is still present in
263 Finland at an estimated prevalence of 1.8% [32]. Fortunately, the deficit in the clinical test has
264 been recognized and remediated by the manufacturer via incorporation of an additional
265 chlamydial target [26] to improve reliability.

266 While discordancy between the clinical diagnostic assay for CT 23S RNA and detection
267 of CT 16S rRNA, whether by qPCR or in the 16S rRNA libraries generated for microbiome analysis
268 was ultimately elucidated by our identification of a genetic variant, the basis for the discrepancy
269 of detection of NG 16S ASVs in participant samples from study participants who were diagnosed
270 NG negative via the AC2 assay is not easily explained. The NG probe of the AC2 assay is directed
271 against 16S rRNA and is both sensitive and highly specific [33-35]. Multiplex diagnostic tests have
272 been reported to have high sensitivity and high concordance with standard nucleic acid
273 amplification testing [36, 37] with a 4-in-1 test for CT, NG, *Trichomonas vaginalis*, and *M.*
274 *genitalium* receiving recent clearance by the FDA [38]. False positive results for NG have been
275 obtained arising from recombination between 16S loci of *N. meningitidis*, now recognized as a
276 urogenital pathogen [39-42] and NG [43, 44]. The presence of multiple ribosomal RNA operons
277 in this genus [45] elevates the risk for even rare recombination events to become fixed in the
278 population. Should the inverse recombination event also occur, and NG acquire 16S copies from
279 other neisserial species, this could lead to reduced sensitivity of currently available diagnostic
280 tests, the consequences of which could negatively impact infection control and promote spread

281 of diagnostic-avoidant mutants as was observed with Swedish nvCT [46, 47]. Identification of
282 highly conserved, alternative targets for detection of NG has proven extremely challenging, so it
283 is not surprising that we were unable to positively confirm the presence of genomic DNA in all the
284 samples. However, we were also unable to amplify even multicopy targets from specimens
285 collected from participants that tested positively by clinical NAAT, raising the possibility of low
286 abundance bacteria that are asymmetrically distributed between the cervix and vaginal vault of
287 discrepant enrollees.

288 Fortunately, study participants of TRAC were not harmed by failure to detect their CT
289 and/or NG infection because all recipients were administered appropriate antibiotic therapy at
290 enrollment regardless of their NAAT result. Appropriately, health care providers are moving from
291 empiric syndromic treatment of patients at risk for STIs to pathogen-specific treatment guided by
292 results of molecular tests [48]. Untreated CT and NG infections can lead to severe complications
293 including PID, infertility, ectopic pregnancy and chronic pelvic pain in women. While NAATs are
294 highly accurate, our study demonstrates that clinicians and epidemiologists should remain vigilant
295 for potential CT or NG variants evading current approved NAATs such as detection of negative
296 test results in a recently exposed individual, or a decline in overall rates of infection in a particular
297 geographic location as occurred with Swedish nvCT [46, 47].

298

299 **Declarations:**

300 **Ethics approval and consent to participate:** The study protocol was approved by the
301 institutional review boards for human research of the University of Pittsburgh (#PRO10010159)
302 and the University of North Carolina (#13-3074). All participants provided written informed consent
303 at the time of enrollment.

304 **Consent for publication:** Not applicable.

Availability of data and materials: The raw sequencing data supporting the conclusions of this article is publicly available at NCBI's Short Read Archive (SRA) under BioProject accession number PRJNA1136868.

Competing interests: The authors declare that they have no competing interests.

Funding: This work was supported by the National Institute of Allergy and Infectious Diseases via U19 AI084024 and R01 AI170959.

Authors' contributions: SJ, KSY, XP, and CMO designed research studies. HCW and SLH directed collection of participant samples and performance of diagnostic testing. TT, HC, TY and PB performed experiments. SJ, TT, HB, HC, and XP analyzed data. SJ, KSY, TD, XP, and CMO wrote the manuscript. All authors reviewed the manuscript.

Acknowledgements: We thank the women who agreed to participate in this study; Ingrid Macio, Melinda Petrina, Carol Priest, Abi Jett, and Lorna Rabe for their efforts in the clinic and the microbiology laboratory; and the staff at the Allegheny County Health Department STD Clinic, for their efforts. We would also like to acknowledge the contribution of Dr. Ian Huntress, who sadly passed away during the performance of these studies.

321 **References:**

322 1. Russell AN, Zheng X, O'Connell CM, Taylor BD, Wiesenfeld HC, Hillier SL, Zhong W,
323 Darville T: **Analysis of Factors Driving Incident and Ascending Infection and the**
324 **Role of Serum Antibody in Chlamydia trachomatis Genital Tract Infection**. *J Infect*
325 *Dis* 2016, **213**(4):523-531.
326 2. Russell AN, Zheng X, O'Connell CM, Wiesenfeld HC, Hillier SL, Taylor BD, Picard MD,
327 Flechtner JB, Zhong W, Frazer LC *et al*: **Identification of Chlamydia trachomatis**
328 **Antigens Recognized by T Cells From Highly Exposed Women Who Limit or**
329 **Resist Genital Tract Infection**. *J Infect Dis* 2016, **214**(12):1884-1892.
330 3. Liu C, Mokashi Neha V, Darville T, Sun X, O'Connell Catherine M, Hufnagel K,
331 Waterboer T, Zheng X: **A Machine Learning-Based Analytic Pipeline Applied to**
332 **Clinical and Serum IgG Immunoproteome Data To Predict Chlamydia trachomatis**
333 **Genital Tract Ascension and Incident Infection in Women**. *Microbiology Spectrum*
334 2023, **0**(0):e04689-04622.
335 4. Liu C, Hufnagel K, O'Connell CM, Goonetilleke N, Mokashi N, Waterboer T, Tollison TS,
336 Peng X, Wiesenfeld HC, Hillier SL *et al*: **Reduced Endometrial Ascension and**
337 **Enhanced Reinfection Associated With Immunoglobulin G Antibodies to Specific**
338 **Chlamydia trachomatis Proteins in Women at Risk for Chlamydia**. *J Infect Dis* 2022,
339 **225**(5):846-855.
340 5. Darville T, Albritton HL, Zhong W, Dong L, O'Connell CM, Poston TB, Quayle AJ,
341 Goonetilleke N, Wiesenfeld HC, Hillier SL *et al*: **Anti-chlamydia IgG and IgA are**
342 **insufficient to prevent endometrial chlamydia infection in women, and increased**
343 **anti-chlamydia IgG is associated with enhanced risk for incident infection**. *Am J*
344 *Reprod Immunol* 2019, **81**(5):e13103.

6. Poston TB, Lee DE, Darville T, Zhong W, Dong L, O'Connell CM, Wiesenfeld HC, Hillier SL, Sempowski GD, Zheng X: **Cervical Cytokines Associated With Chlamydia trachomatis Susceptibility and Protection**. *J Infect Dis* 2019, **220**(2):330-339.

7. Zheng X, O'Connell CM, Zhong W, Nagarajan UM, Tripathy M, Lee D, Russell AN, Wiesenfeld H, Hillier S, Darville T: **Discovery of Blood Transcriptional Endotypes in Women with Pelvic Inflammatory Disease**. *J Immunol* 2018, **200**(8):2941-2956.

8. Zheng X, Zhong W, O'Connell CM, Liu Y, Haggerty CL, Geisler WM, Anyalechi GE, Kirkcaldy RD, Wiesenfeld HC, Hillier SL *et al*: **Host Genetic Risk Factors for Chlamydia trachomatis-Related Infertility in Women**. *J Infect Dis* 2021, **224**(Supplement_2):S64-s71.

9. Zheng X, O'Connell CM, Zhong W, Poston TB, Wiesenfeld HC, Hillier SL, Trent M, Gaydos C, Tseng G, Taylor BD *et al*: **Gene Expression Signatures Can Aid Diagnosis of Sexually Transmitted Infection-Induced Endometritis in Women**. *Front Cell Infect Microbiol* 2018, **8**(307):307.

10. Taylor BD, Zheng X, O'Connell CM, Wiesenfeld HC, Hillier SL, Darville T: **Risk factors for Mycoplasma genitalium endometritis and incident infection: a secondary data analysis of the T cell Response Against Chlamydia (TRAC) Study**. *Sex Transm Infect* 2018, **94**(6):414-420.

11. Jeong S, Tollison T, Brochu H, Chou H, Huntress I, Yount KS, Zheng X, Darville T, O’Connell CM, Peng X: **Cervicovaginal microbial features predicts Chlamydia trachomatis spread to the upper genital tract of infected women**. *bioRxiv* 2024:2024.2011.2026.625070.

12. Cole MJ, Davis GS, Fifer H, Saunders JM, Unemo M, Hadad R, Roberts DJ, Fazal M, Day MJ, Minshull J *et al*: **No widespread dissemination of Chlamydia trachomatis diagnostic-escape variants and the impact of Neisseria gonorrhoeae positivity on the Aptima Combo 2 assay**. *Sex Transm Infect* 2022, **98**(5):366-370.

371 13. Hadad R, Jensen JS, Westh H, Gronbaek I, Schwartz LJ, Nielsen L, Vang TM, Nielsen
372 R, Weinreich LS, Skov MN *et al*: **A Chlamydia trachomatis 23S rRNA G1523A variant**
373 **escaping detection in the Aptima Combo 2 assay (Hologic) was widespread**
374 **across Denmark in July-September 2019**. *APMIS* 2020, **128**(6):440-444.
375 14. Roberts DJ, Davis GS, Cole MJ, Naik D, Maru H, Woodford N, Muir P, Horner P, Simms
376 I, Thickett G *et al*: **Prevalence of new variants of Chlamydia trachomatis escaping**
377 **detection by the Aptima Combo 2 assay, England, June to August 2019**. *Euro*
378 *Surveill* 2019, **24**(38).
379 15. Edgar RC: **Search and clustering orders of magnitude faster than BLAST**.
380 *Bioinformatics* 2010, **26**(19):2460-2461.
381 16. O'Connell CM, Brochu H, Girardi J, Harrell E, Jones A, Darville T, Sena AC, Peng X:
382 **Simultaneous profiling of sexually transmitted bacterial pathogens, microbiome,**
383 **and concordant host response in cervical samples using whole transcriptome**
384 **sequencing analysis**. *Microb Cell* 2019, **6**(3):177-183.
385 17. Hazel A, Ponnaluri-Wears S, Davis GS, Low BS, Foxman B: **High prevalence of**
386 **Neisseria gonorrhoeae in a remote, undertreated population of Namibian**
387 **pastoralists**. *Epidemiol Infect* 2014, **142**(11):2422-2432.
388 18. Mitchev N, Singh R, Ramsuran V, Ismail A, Allam M, Kwenda S, Mnyameni F, Garrett N,
389 Swe Swe-Han K, Niehaus AJ *et al*: **Assessment of Antibiotic Resistance and Efflux**
390 **Pump Gene Expression in Neisseria Gonorrhoeae Isolates from South Africa by**
391 **Quantitative Real-Time PCR and Regression Analysis**. *Int J Microbiol* 2022,
392 **2022**:7318325.
393 19. Hokynar K, Rantakokko-Jalava K, Hakanen A, Havana M, Mannonen L, Jokela P,
394 Kurkela S, Lappalainen M, Unemo M, Puolakkainen M: **The Finnish new variant of**
395 **Chlamydia trachomatis with a single nucleotide polymorphism in the 23S rRNA**

396 **target escapes detection by the Aptima Combo 2 test**. *Microorganisms* 2019,
397 **7**(8):227.
398 20. **R: A Language and Environment for Statistical Computing**. In. Edited by Team RC:
399 R Foundation for Statistical Computing; 2015.
400 21. Wickham H: **ggplot: Elegant graphics for data analysis**. New York: Springer-Verlag
401 2016.
402 22. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A,
403 Kuske CR, Tiedje JM: **Ribosomal Database Project: data and tools for high**
404 **throughput rRNA analysis**. *Nucleic Acids Res* 2014, **42**(Database issue):D633-642.
405 23. Robert CE: **SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS**
406 **sequences**. *bioRxiv* 2016:074161.
407 24. Cook RL, Hutchison SL, Ostergaard L, Braithwaite RS, Ness RB: **Systematic review:**
408 **noninvasive testing for Chlamydia trachomatis and Neisseria gonorrhoeae**. *Ann*
409 *Intern Med* 2005, **142**(11):914-925.
410 25. Centers for Disease C, Prevention: **Recommendations for the laboratory-based**
411 **detection of Chlamydia trachomatis and Neisseria gonorrhoeae--2014**. *MMWR*
412 *Recomm Rep* 2014, **63**(RR-02):1-19.
413 26. Unemo M, Hansen M, Hadad R, Puolakkainen M, Westh H, Rantakokko-Jalava K,
414 Thilesen C, Cole MJ, Boiko I, Lan PT *et al*: **Sensitivity, specificity, inclusivity and**
415 **exclusivity of the updated Aptima Combo 2 assay, which provides detection**
416 **coverage of the new diagnostic-escape Chlamydia trachomatis variants**. *BMC*
417 *Infect Dis* 2020, **20**(1):419.
418 27. Rantakokko-Jalava K, Hokynar K, Hieta N, Keskitalo A, Jokela P, Muotiala A, Jokiranta
419 TS, Kuusela R, Sarkkinen H, Aittoniemi J *et al*: **Chlamydia trachomatis samples**
420 **testing falsely negative in the Aptima Combo 2 test in Finland, 2019**. *Euro Surveill*
421 2019, **24**(22).

422 28. Smith NH, Holmes EC, Donovan GM, Carpenter GA, Spratt BG: **Networks and groups**
423 **within the genus Neisseria: analysis of argF, recA, rho, and 16S rRNA sequences**
424 **from human Neisseria species**. *Mol Biol Evol* 1999, **16**(6):773-783.
425 29. Abrams AJ, Trees DL, Nicholas RA: **Complete Genome Sequences of Three**
426 **Neisseria gonorrhoeae Laboratory Reference Strains, Determined Using PacBio**
427 **Single-Molecule Real-Time Technology**. *Genome Announc* 2015, **3**(5).
428 30. Ramsden S, Ovens K, Griffiths J, Muir P, Steele-Nicholson A, Horner PJ:
429 **Environmental contamination by Chlamydia trachomatis and Neisseria**
430 **gonorrhoeae: is it time to change our infection control practices? Results of a**
431 **regional study**. *Sex Transm Infect* 2023, **99**(1):35-40.
432 31. Katz SS, Danavall DC, Morris MR, Herrod BP, Dale SE, Nye MB, Kersh EN, Kirkcaldy
433 RD, Raphael BH: **Chlamydia trachomatis Variants Escaping Detection in the**
434 **Aptima Combo 2 Assay in the United States**. *Sex Transm Dis* 2022, **49**(6):448-452.
435 32. Puolakkainen M, Mannonen L, Unemo M: **The diagnostic-escape Finnish new variant**
436 **of Chlamydia trachomatis first described in 2019 remains present in Finland three**
437 **years after the introduction of an updated Aptima Combo 2 assay**. *Clin Microbiol*
438 *Infect* 2023, **29**(5):658-659.
439 33. Bristow CC, McGrath MR, Cohen AC, Anderson LJ, Gordon KK, Klausner JD:
440 **Comparative evaluation of two nucleic acid amplification tests for the detection of**
441 **Chlamydia trachomatis and Neisseria gonorrhoeae at extragenital sites**. *Sexually*
442 *transmitted diseases* 2017, **44**(7):398.
443 34. Golparian D, Tabrizi SN, Unemo M: **Analytical specificity and sensitivity of the**
444 **APTIMA Combo 2 and APTIMA GC assays for detection of commensal Neisseria**
445 **species and Neisseria gonorrhoeae on the Gen-Probe Panther instrument**.
446 *Sexually Transmitted Diseases* 2013, **40**(2):175-178.

447 35. Tabrizi SN, Unemo M, Limnios AE, Hogan TR, Hjelmevoll S-O, Garland SM, Tapsall J:
448 **Evaluation of six commercial nucleic acid amplification tests for detection of**
449 **Neisseria gonorrhoeae and other Neisseria species**. *Journal of clinical microbiology*
450 2011, **49**(10):3610-3615.
451 36. Matysiak C, Cheng A, Kirby JE: **Evaluation of the Abbott Alinity m STI assay for**
452 **diagnosis of the primary causes of sexually transmitted infections in the United**
453 **States**. *Pract Lab Med* 2023, **36**:e00332.
454 37. Kriesel JD, Bhatia AS, Barrus C, Vaughn M, Gardner J, Crisp RJ: **Multiplex PCR**
455 **testing for nine different sexually transmitted infections**. *Int J STD AIDS* 2016,
456 **27**(14):1275-1282.
457 38. Administration USFD: **Alinity m CMV Assay - P210022**. In.; 2022.
458 39. Urra E, Alkorta M, Sota M, Alcalá B, Martínez I, Barrón J, Cisterna R: **Orogenital**
459 **transmission of Neisseria meningitidis serogroup C confirmed by genotyping**
460 **techniques**. *Eur J Clin Microbiol Infect Dis* 2005, **24**(1):51-53.
461 40. Katz AR, Chasnoff R, Komeya A, Lee MV: **Neisseria meningitidis urethritis: a case**
462 **report highlighting clinical similarities to and epidemiological differences from**
463 **gonococcal urethritis**. *Sex Transm Dis* 2011, **38**(5):439-441.
464 41. Kanemitsu N, Hayashi I, Satoh N, Hayakawa T, Mitsuya H, Hayase Y, Hiramoto K,
465 Kojima M: **Acute urethritis caused by Neisseria meningitidis**. *Int J Urol* 2003,
466 **10**(6):346-347.
467 42. Bazan JA, Peterson AS, Kirkcaldy RD, Briere EC, Maierhofer C, Turner AN, Licon DB,
468 Parker N, Dennison A, Ervin M *et al*: **Notes from the Field: Increase in Neisseria**
469 **meningitidis-Associated Urethritis Among Men at Two Sentinel Clinics -**
470 **Columbus, Ohio, and Oakland County, Michigan, 2015**. *MMWR Morb Mortal Wkly*
471 *Rep* 2016, **65**(21):550-552.

472 43. Walcher M, Skvoretz R, Montgomery-Fullerton M, Jonas V, Brentano S: **Description of**
473 **an unusual Neisseria meningitidis isolate containing and expressing Neisseria**
474 **gonorrhoeae-Specific 16S rRNA gene sequences**. *J Clin Microbiol* 2013,
475 **51**(10):3199-3206.
476 44. Sukhum KV, Jean S, Wallace M, Anderson N, Burnham CA, Dantas G: **Genomic**
477 **Characterization of Emerging Bacterial Uropathogen Neisseria meningitidis,**
478 **Which Was Misidentified as Neisseria gonorrhoeae by Nucleic Acid Amplification**
479 **Testing**. *J Clin Microbiol* 2021, **59**(2).
480 45. Klappenbach JA, Saxman PR, Cole JR, Schmidt TM: **rrndb: the Ribosomal RNA**
481 **Operon Copy Number Database**. *Nucleic Acids Res* 2001, **29**(1):181-184.
482 46. Ripa T, Nilsson P: **A variant of Chlamydia trachomatis with deletion in cryptic**
483 **plasmid: implications for use of PCR diagnostic tests**. *EuroSurveill* 2006,
484 **11**(11):E061109.
485 47. Unemo M, Seth-Smith HM, Cutcliffe LT, Skilton RJ, Barlow D, Goulding D, Persson K,
486 Harris SR, Kelly A, Bjartling C *et al*: **The Swedish new variant of Chlamydia**
487 **trachomatis: genome sequence, morphology, cell tropism and phenotypic**
488 **characterization**. *Microbiology*, **156**(Pt 5):1394-1404.
489 48. Karellis A, Naeem F, Nair S, Mallya SD, Routy JP, Gahagan J, Yansouni CP, Kim J, Pai
490 NP: **Multiplexed rapid technologies for sexually transmitted infections: a**
491 **systematic review**. *Lancet Microbe* 2022, **3**(4):e303-e315.
492
493

Table 2. *Neisseria sp.* transcripts detected by 16S ASV analysis and genes detected by qPCR in diagnostic-discrepant or NG+ TRAC participants.

Sample ID	Diagnostic test		16S rRNA read counts				qPCR†		
	Cervix	Endo	ZOtu95: N_gonorrhoeae	ZOtu141: N_subflava	ZOtu264: N_mucosa	ZOtu1566: N_elongata	16S	23S	ParA
3-121	-	-	130	0	0	0	+	+	-
3-140	-	-	123	0	0	0	+	+	+
3-146	-	-	106	0	0	0	+	+	ND
3-151	-	-	6	0	0	0	+	-	+
3-209	-	-	1	0	0	0	-	+	ND
3-012	+	-	471	0	0	0	+	+	+
3-026	+	-	10	0	0	0	-	-	-
3-052	+	-	10	0	0	0	-	+	-
3-089	+	-	24	0	0	0	-	+	+
3-091	+	-	6	0	0	0	+	+	+
3-112	+	-	1032	0	0	0	+	+	+
3-114	+	-	85	0	0	0	-	-	-
3-120	+	-	19	211	131	0	-	+	-
3-122	+	-	110	0	0	0	+	+	+
3-124	+	-	927	0	0	0	-	+	-
3-133	+	-	3	0	0	0	-	-	-
3-136	+	-	2	0	1	0	+	+	-
3-177	+	-	180	0	0	0	+	+	+
3-208	+	-	1151	0	0	0	-	+	-
3-212	+	-	9	0	0	0	+	+	+
3-223	+	-	206	0	0	0	+	+	+
3-235	+	-	349	0	0	0	+	+	+

†Assay performed with diagnostic swab eluate or vaginal swab template used for 16S rRNA microbiome analysis. ND not done because specimen was insufficient or unavailable.

494

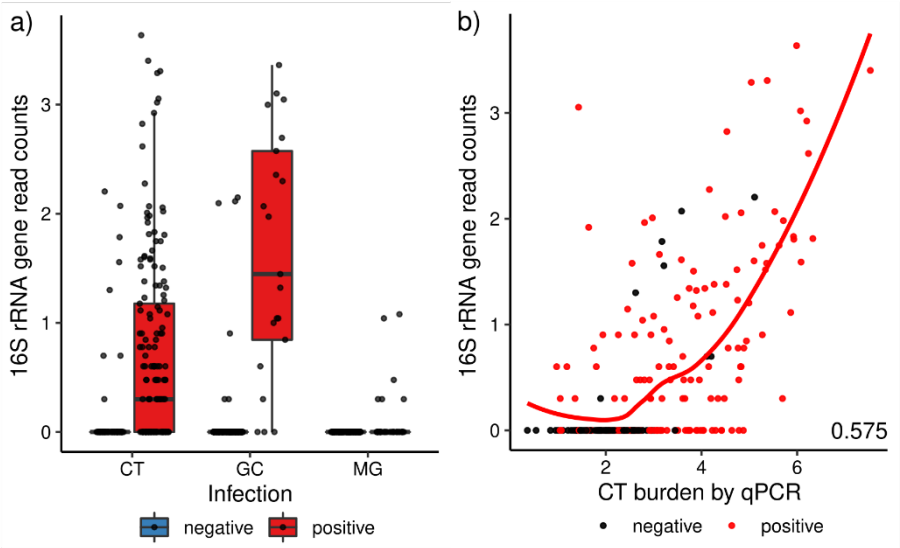
495

496

497

498

499



500

501 **Figure 1. STI-derived 16S transcripts detected in vaginal swabs**. a) 16S rRNA gene read
502 counts of STIs stratified by infection status at enrollment. b) Concordance between CT 16S rRNA
503 gene read counts versus CT burden by qPCR, where samples are colored by infection status at
504 enrollment. A LOESS curve is shown in red and Spearman correlation in the bottom-right. CT,
505 *Chlamydia trachomatis*; GC, *N. gonorrhoeae*; MG, *M. genitalium*

506

507

```
                1450      1460      1470      1480      1490      1500      1510      1520      1530      1540      1550      1560      1570      1580
                   |         |         |         |         |         |         |         |         |         |         |         |         |         |
Reference  1446 gtccgtagagcgatgagaacggttagtaggcaaatccgctaacataagatcaggtcgcgatcaaggggaatcttcgggggaaccgatggtgtggagcgaggctttcaagaaataatttctagctgttgatggtgaccgt
3_144        21 gtccgtagagcgatgagaacggttagtaggcaaatccgctaacataagatcaggtcgcgatcaaggggaatcttcggaggaaccgatggtgtggagcgaggctttcaagaaataatttctagctgttgatggtgaccgt
3_145        22 gtccgtagagcgatgagaacggttagtaggcaaatccgctaacataagatcaggtcgcgatcaaggggaatcttcggaggaaccgatggtgtggagcgaggctttcaagaaataatttctagctgttgatggtgaccgt
3_148        23 gtccgtagagcgatgagaacggttagtaggcaaatccgctaacataagatcaggtcgcgatcaaggggaatcttcggaggaaccgatggtgtggagcgaggctttcaagaaataatttctagctgttgatggtgaccgt
3_226        22 gtccgtagagcgatgagaacggttagtaggcaaatccgctaacataagatcaggtcgcgatcaaggggaatcttcggaggaaccgatggtgtggagcgaggctttcaagaaataatttctagctgttgatggtgaccgt
3_209        22 gtccgtagagcgatgagaacggttagtaggcaaatccgctaacataagatcaggtcgcgatcaaggggaatcttcggaggaaccgatggtgtggagcgaggctttcaagaaataatttctagctgttgatggtgaccgt
3_101        22 gtccgtagagcgatgagaacggttagtaggcaaatccgctaacataagatcaggtcgcgatcaaggggaatcttcggaggaaccgatggtgtggagcgaggctttcaagaaataatttctagctgttgatggtgaccgt
3_201        22 gtccgtagagcgatgagaacggttagtaggcaaatccgctaacataagatcaggtcgcgatcaaggggaatcttcggaggaaccgatggtgtggagcgaggctttcaagaaataatttctagctgttgatggtgaccgt
```

508

509 **Figure 2. Portion of sequence alignment of V3 amplified PCR products from discrepant**

510 **TRAC participant specimens revealing a g to a substitution at 1523 in the 23S rDNA locus**.

511 PCR products from two representative CT+ samples, as determined by AC2 diagnostic test were

512 wild type at this position (data not shown).

513
514