

OPEN

The association of *Chlamydia trachomatis* and *Mycoplasma genitalium* infection with the vaginal metabolome

Joanna-Lynn C. Borgogna¹, Michelle D. Shardell^{2,3}, Carl J. Yeoman^{1,4}, Khalil G. Ghanem⁵, Herlin Kadriu¹, Alexander V. Ulanov⁶, Charlotte A. Gaydos⁵, Justin Hardick⁵, Courtney K. Robinson³, Patrik M. Bavoil⁷, Jacques Ravel^{3,8}, Rebecca M. Brotman^{2,3} & Susan Tuddenham^{5*}

Chlamydia trachomatis (CT) and *Mycoplasma genitalium* (MG) are two highly prevalent bacterial sexually transmitted infections (STIs) with a significant rate of co-infection in some populations. Vaginal metabolites are influenced by resident vaginal microbiota, affect susceptibility to sexually transmitted infections (STIs), and may impact local inflammation and patient symptoms. Examining the vaginal metabolome in the context of CT mono (CT+) and CT/MG co-infection (CT+/MG+) may identify biomarkers for infection or provide new insights into disease etiology and pathogenesis. Yet, the vaginal metabolome in the setting of CT infection is understudied and the composition of the vaginal metabolome in CT/MG co-infected women is unknown. Therefore, in this analysis, we used an untargeted metabolomic approach combined with 16S rRNA gene amplicon sequencing to characterize the vaginal microbiota and metabolomes of CT+, CT+/MG+, and uninfected women. We found that CT+ and CT+/MG+ women had distinct vaginal metabolomic profiles as compared to uninfected women both before and after adjustment for the vaginal microbiota. This study provides important foundational data documenting differences in the vaginal metabolome between CT+, CT+/MG+ and uninfected women. These data may guide future mechanistic studies that seek to provide insight into the pathogenesis of CT and CT/MG infections.

Vaginal metabolites are influenced by resident vaginal microbiota, affect susceptibility to sexually transmitted infections (STIs), and impact local inflammation and symptoms^{1,2}. An “optimal” vaginal microbiota is dominated by *Lactobacillus* species, which produce the metabolite lactic acid. Lactic acid lowers the pH of the vaginal microenvironment^{1–4}, and, through immunomodulatory and direct inhibitory effects, may protect against acquisition of STIs, including *Chlamydia trachomatis* (CT) and HIV^{2,5–7}. In contrast, women with non-optimal microbiota, as epitomized by the clinical condition of bacterial vaginosis (BV), have vaginal microbial communities that are low in *Lactobacillus* spp.; and are instead colonized by a variety of anaerobes which generally produce very little or no lactic acid. Some of these bacteria produce metabolites such as biogenic amines and short chain fatty acids that may be pro-inflammatory and have been linked with symptoms such as vaginal malodor and discomfort^{2,8–12}. These metabolites may also increase susceptibility to STIs, including bolstering chlamydial persistence^{8,13–16}. Women with low-*Lactobacillus* non-optimal vaginal microbiota have an increased risk for acquisition of STIs and ascending infection (including pelvic inflammatory disease [PID])^{17–21}.

¹Department of Microbiology and Immunology, Montana State University, Bozeman, MT, USA. ²Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, MD, USA. ³Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD, USA. ⁴Department of Animal and Range Sciences, Montana State University, Bozeman, MT, USA. ⁵Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ⁶Roy J. Carver Biotechnology Center, University of Illinois, Urbana-Champaign, IL, USA. ⁷Department of Microbial Pathogenesis, University of Maryland School of Medicine, Baltimore, MD, USA. ⁸Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, MD, USA. *email: studden1@jhmi.edu

CT is the most prevalent bacterial STI in the United States, and causes significant morbidity, including PID and infertility in young women²². *Mycoplasma genitalium* (MG) is an emerging, highly drug resistant STI, which appears to be more prevalent than gonorrhoea, and has been linked to cervicitis and PID in women^{23–25}, though current CDC guidelines recommend consideration for testing and treatment only in women with persistent cervicitis²⁶.

Interestingly, rates of co-infection with CT and MG are high in many female populations (36% of CT infected women had MG co-infection in one study done in Baltimore, MD,²⁷ nearly 38% of CT infected women were found to have MG co-infection in a second study²⁸ and 70.7% of women with MG in a high risk cohort were found to have CT co-infection²⁵. Endometrial MG infection has been associated with endometritis and endometrial CT infection, suggesting that these infections may co-locate or perhaps affect PID risk in some women²⁵; however the impact of co-infection on patient outcomes is understudied¹.

New molecular techniques have enabled a higher resolution understanding of BV, the vaginal microbiota, and associations with STI, including CT and MG infections. 16S rRNA gene amplicon sequencing techniques have identified several different kinds vaginal microbiota, or “community state types” (CSTs)²⁹ based on bacterial composition and relative abundance. Four are dominated by different *Lactobacillus* species (CST I, *Lactobacillus crispatus*; CST II, *L. gasseri*; CST III, *L. iners*; CST V, *L. jensenii*). A fifth community (CST IV) is low in *Lactobacillus* spp. and has been termed “molecular-Bacterial Vaginosis” (molecular-BV)³. CST IV and CST III have been associated with increased risk of CT acquisition³⁰. Emerging evidence suggests that low-*Lactobacillus* vaginal microbiota may be associated with increased risk for MG infection as well^{18,30}.

Associations between the vaginal microbiota and enhanced risk of CT and MG infection are likely mediated, at least in part, through vaginal metabolites. CT and MG may also independently influence the vaginal metabolome through direct, host-mediated or microbiota-mediated mechanisms, perhaps impacting clinical outcomes. Therefore, examining the vaginal metabolome in the context of these infections may identify biomarkers for infection or provide new insights into disease etiology and pathogenesis. Yet, to date, only one study has examined the vaginal metabolome in the setting of CT infection³¹, and the composition of the vaginal metabolome in CT/MG co-infected women is unknown. Therefore, in this study, we utilized untargeted gas chromatography-mass spectrometry (GC-MS) to characterize the vaginal metabolome of CT mono-, CT/MG co-infected, and uninfected women. As STI-infected women are known to have a higher prevalence of BV^{30,32,33}, we additionally integrated analysis of the vaginal metabolome with 16S rRNA gene amplicon sequencing in order to understand whether the identified differences in vaginal metabolites remained statistically significant after controlling for the composition of the vaginal microbiota.

Results

Participant characteristics and vaginal microbiota composition (CST) by infection state.

Vaginal samples from 145 race-matched women (77 uninfected, 54 CT+ (mono-infected), and 14 CT+/MG+ (co-infected)) were analyzed. CT+ and CT+/MG+ women were generally younger, had more sexual partners, and a lower proportion reported “always” using condoms as compared to uninfected women (Supplementary Table S1). Seven (50%) of the CT+/MG+ women had evidence of macrolide resistance based on the SpeedX MG testing. A much larger proportion of CT+ (64.8%) and CT+/MG+ (78.6%) women had CST IV (low-*Lactobacillus* species vaginal microbiota) as compared with uninfected women (26.0%, $p < 0.01$), and Supplementary Fig. S1). The majority of women (45 of the 54 CT mono-infected women (83%) and 12 of the 14 CT+/MG+ co-infected women (85%)) reported one or more symptom or were diagnosed with pelvic inflammatory disease on exam.

Vaginal metabolomic profiles distinguish women based on infection status and CST.

An exploratory PCA plot revealed that metabolomes clustered by infection status with infected women clustering separately from uninfected women, and CT+ and CT+/MG+ women largely clustering together (Fig. 1A). A second PCA showed that women in CST IV (low-*Lactobacillus* vaginal microbiota) clustered separately from those with *L. crispatus*-, *L. gasseri*-, or *L. jensenii*- dominated vaginal microbiota (CST I, II, and V, respectively). Women with *L. iners* dominated vaginal microbiota (CST III) were interspersed between the two clusters (Fig. 1B). Both multiple linear regression and PLS-DA identified several metabolites that significantly distinguished: (1) CT+ vs. uninfected women; (2) CT+/MG+ vs. uninfected women; and (3) CT+/MG+ vs. CT+ women (Supplementary Tables S2 and S3). Given the demographic and behavioral differences between cases and controls (Supplementary Table S1), we also assessed the potential impact of these factors upon the vaginal metabolome. Results from this analysis indicate that age, number of recent sexual partners, and condom use, among others, were not significantly associated with the vaginal metabolome. (See Supplementary Table S4). The only significant variables that contributed >3% of the variation to the vaginal metabolome were infection status (12%) and CST (10%).

Multiple vaginal metabolites distinguish CT mono-, CT/MG co-infected and uninfected women before and after adjustment for CST.

Multiple linear regression demonstrated that there were significant differences in individual vaginal metabolites in CT+ and CT+/MG+ compared with uninfected women both before and after adjustment for CST. Supplementary Tables S2 and S3 summarize all significant metabolites identified by multiple linear regression, grouped by metabolic pathway based on the Kyoto Encyclopedia of Genes and Genomes (KEGG). All results reported below have a q -value < 0.05 unless otherwise stated.

In unadjusted analyses, the relative concentrations of 99 and 64 metabolites were significantly different when comparing CT+ and CT+/MG+ women to uninfected women, respectively (Supplementary Table S2). After adjustment for CST, 64 and 36 metabolites significantly differed when comparing CT+ and CT+/MG+ co-infected women, respectively, to uninfected women (Supplementary Table S3). 93% of metabolites CT+ vs. uninfected patients (60/64) and 94% of metabolites in CT+/MG+ co-infected women vs. uninfected women

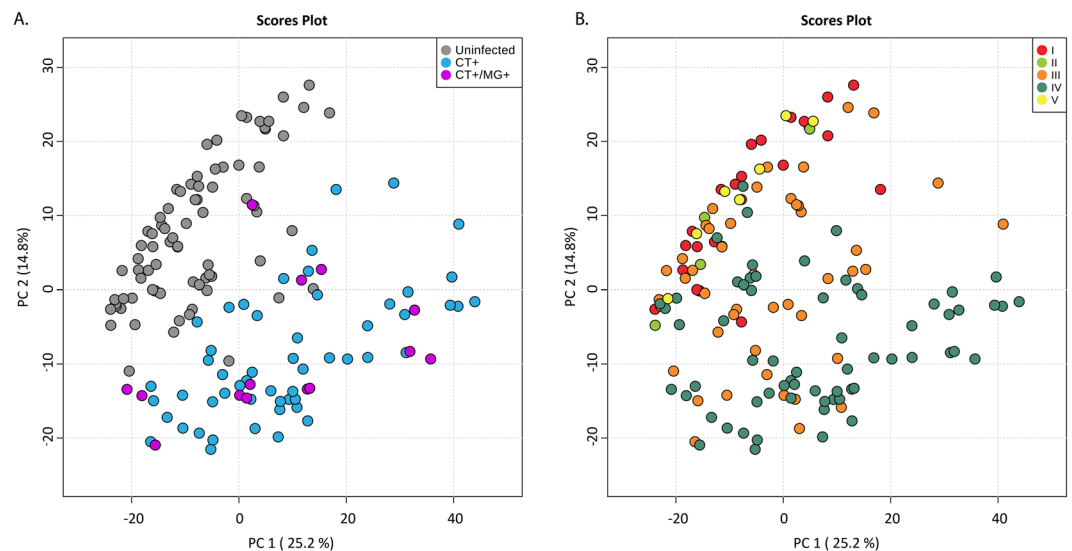


Figure 1. The vaginal metabolome clusters by infection status and by CST. Two-dimensional principal component analysis (PCA) score plots demonstrate statistical clustering of metabolites by (A) infection status and (B) by CST.

(34/36) which were significant before adjustment for CST were also significant after adjustment for CST. Comparatively fewer metabolites differentiated CT+/MG+ and CT+ women both before (7 metabolites) and after adjustment for CST (3 metabolites). All of the metabolites that distinguished these two groups after adjustment for CST were significant prior to adjustment for CST.

Amino acid metabolism. Women with CST IV irrespective of infection status, had greater vaginal concentrations of the biogenic amines cadaverine, putrescine, tyramine and ethanolamine (part of the lipid super pathway but is also considered a biogenic amine) based on linear regression (one-sided p -value < 0.05) compared with women with *Lactobacillus*-dominated CSTs. Consistent with this result, (given the higher proportion of women with CST IV microbiota in the infected groups), before adjustment for CST, the biogenic amines cadaverine, putrescine, tyramine and ethanolamine were increased in the CT+, and cadaverine, putrescine and ethanolamine were increased in the CT+/MG+ as compared with the uninfected women. After adjustment for CST, only ethanolamine was increased in infected as compared with uninfected patients.

Cinnamic acid or its associated derivatives were higher in CT+ and CT+/MG+ women compared to uninfected women with and without adjustment for CST. Hydrocinnamic acid was increased in CT+/MG+ as compared to CT+ women, but there were no differences in biogenic amines either before or after adjustment for CST between these two groups.

Carbohydrate metabolism. Several sugars and derivatives, including lactose and sorbitol, were significantly lower in both CT+ and CT+/MG+ women compared to uninfected women, irrespective of adjustment for CST. Metabolites involved in pentose metabolism, such as xylose, were higher in CT+ and CT+/MG+ relative to uninfected women both before and after adjustment for CST. Glyceric acid, which is involved in glycolysis, was decreased in CT+ and CT+/MG+ relative to uninfected women before and after adjustment for CST.

Lipid metabolism. Multiple long chain fatty acids (LCFAs), including decanoic, pentadecanoic, heptadecanoic, oleic, arachidic, behenic, cerotic, and myristic acid were higher in CT+ as compared to uninfected women without adjustment for CST; the majority of these remained significantly higher after adjustment for CST. Fewer LCFAs, including pentadecanoic, heptadecanoic, and oleic acids were elevated in CT+/MG+ as compared to uninfected women, all of which remained significantly higher after adjustment for CST. The LCFA palmitoleic acid was decreased in CT+/MG+ as compared to CT+ women, but only before adjustment for CST. The sterol, 5 α -cholestan-3-ol, was higher in CT mono-infected women compared to uninfected women without adjustment for CST. Multiple long chain fatty alcohols including octadecanol, eicosanol, heptadecanol and hexadecanol, were higher with and without CST adjustment in CT+ and CT+/MG+ as compared with uninfected women. Octadecanol was lower in CT+/MG+ co-infected as compared with CT+ women both before and after adjustment for CST.

Energy metabolism. α -ketoglutaric acid, citric acid and glutaric acid, all metabolites involved in the TCA cycle, were lower in CT+ compared to uninfected women both before and after adjustment for CST. These metabolites were also significantly decreased in CT+/MG+ women relative to uninfected women after adjustment for CST.

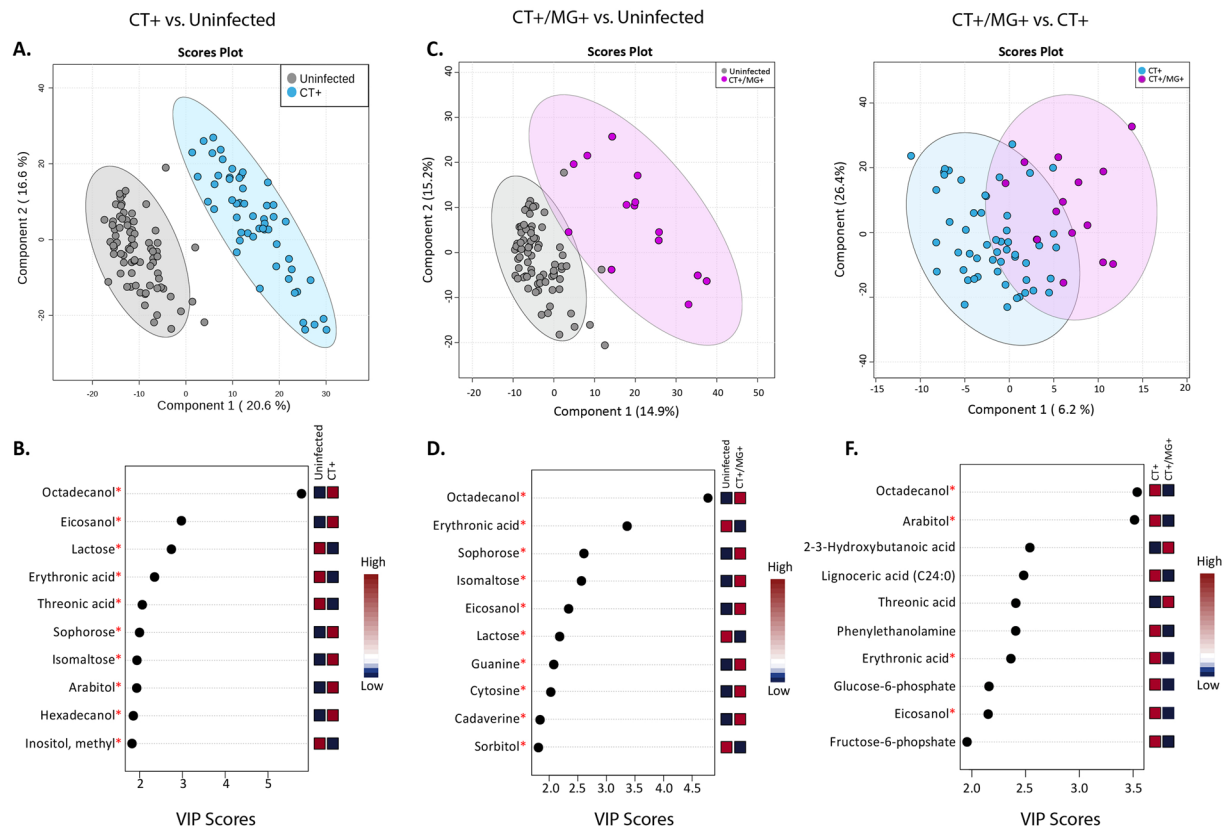


Figure 2. Discriminatory metabolites associated with infection status. CT+ and uninfected women are compared in the first column (A,B), CT+/MG+ and uninfected women are compared in second column (C,D) and CT+/MG+ women compared to CT+ women are compared in the last row (E,F). (A,C,E) Show discrimination of groups using partial least squares discriminant analysis. (B,D,F) show the metabolites most strongly influencing discrimination by the PLS-DA. The variable importance in projection (VIP) score is the weighted sum of squares for the PLS-DA loading with the amount of variation explained by each component taken into account. Asterisks indicate metabolites which are also significant via regression analyses.

Nucleic acid metabolism. Several metabolites involved in purine and pyrimidine metabolism, including adenine, guanine and cytosine, were increased in CT+ relative to uninfected women both before and after adjustment for CST. These were also higher in CT/MG co-infected women before adjustment for CST; guanine and cytosine were higher after adjustment for CST. When comparing CT+/MG+ to CT+ women, we did not observe any significant differences in nucleotide abundance with or without adjustment for CST.

Partial least squares discriminate analysis (PLS-DA). We utilized PLS-DA to identify the most discriminatory metabolites between each infection state (Fig. 2). PLS-DA was able to effectively separate CT+ and CT+/MG+ from uninfected women in two components with accuracies ranging from 94–98%. Permutation tests provided evidence that the differences between these clusters were significant (CT+ vs. uninfected women, p -value < 0.0001 , 0/2000 permutations; CT+/MG+ vs. uninfected women, p -value = 0.0025, 5/2000 permutations). Variable importance plots (VIP) indicated that octadecanol, eicosanol, and lactose (VIP scores > 2.5) contributed the most to the class separation between CT+ and uninfected women, with octadecanol and eicosanol being higher in CT+ women and lactose being lower in CT+ than uninfected women. (Fig. 2B). The most important discriminatory metabolites between CT+/MG+ and uninfected women were octadecanol, erythronic acid, sophorose, and isomaltose with octadecanol, sophorose and isomaltose being higher in CT+/MG+ than uninfected women. (Fig. 2D). PLS-DA was not able to effectively separate CT+/MG+ co-infected women from CT+ women (Fig. 2E). The lack of separation was cross-validated with R^2 coefficients of 0.39 and 0.79 and Q^2 values of -0.35 and 0.22. Permutation tests further provided evidence against the separation of these two groups (p -value 0.167, 334/2000 permutations).

Finally, we subset the data into CST III (*L. iners* dominated) and CST-IV (low-*Lactobacillus*) and performed separate PLS-DAs to evaluate differences between infection status within strata of CST (Fig. 3). CT+ and CT+/MG+ women clustered together and were discriminated from uninfected women, irrespective of CST (CST III, p -value = < 0.0001 , 0/2000 permutations; CST IV p -value < 0.0001 , 0/2000 permutations). Octadecanol, eicosanol and lactose were significant in discriminating between the CT+ and uninfected subjects for women in both CST III and CST IV.

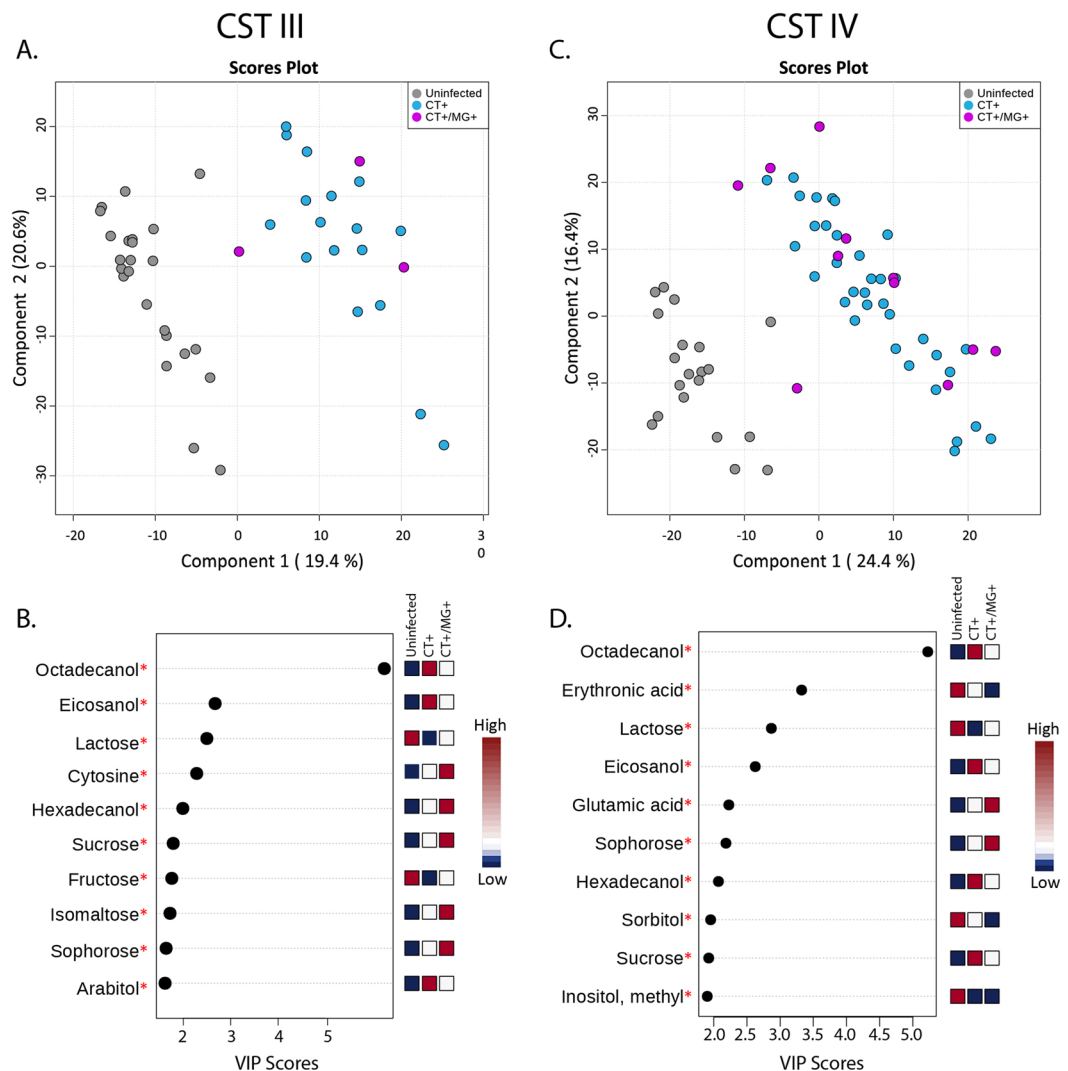


Figure 3. Discriminatory metabolites associated with infection status stratified by CST. Women within CST-III are compared in the first column (A,B) and women within CST-IV are compared in the second column (C,D). (A,C) Show discrimination of groups using partial least squares discriminant analysis. (B,D) Show the metabolites most strongly influencing discrimination by the PLS-DA. Asterisks indicate metabolites which are also significant via regression analyses.

Discussion

We found that numerous vaginal metabolites distinguished CT+ and CT+/MG+ from uninfected women, with and without adjustment for the composition of the vaginal microbiota. When comparing CT+/MG+ to CT+ women, we observed significant overlap of vaginal metabolome profiles.

Previous literature suggests that women with BV are at increased risk for acquisition of CT and MG¹, and women with CT infection are most likely to be in CST IV³⁰. Consistent with this, a larger proportion of the CT+ and CT+/MG+ as compared with the uninfected women in our study cohort had CST IV microbiota. Since biogenic amines (thought to be produced by BV-associated bacteria) are known to be higher in women with CST IV⁸, it is not surprising, therefore, that the biogenic amines putrescine and cadaverine were higher in CT mono- and CT/MG co-infected as compared to uninfected women prior to, but not after adjustment for CST. Biogenic amine production involves specific amino acid decarboxylation reactions, and is a well-described bacterial acid resistance and mitigation mechanism^{8,16,32,34,35}. This reduction of vaginal acidity may promote colonization by other BV-associated bacteria or further overgrowth of existing BV-associated bacteria. Additionally, the production of biogenic amines may increase the virulence of various pathogens, and may protect pathogens from innate immune defenses^{36,37}. Biogenic amines may reduce the host's immune defenses against STIs, including CT or MG⁸.

While some of the differences in vaginal metabolites between infection groups (such as putrescine and cadaverine, above) may relate to the higher proportion of infected women with CST IV, we cannot discount the possibility that these CST-related metabolites are important to the infection process. Differences in other metabolites remained consistent even after adjustment for CST, suggesting a role for the infecting pathogen(s) (or perhaps an interaction between the host, pathogen and underlying microbiota) in impacting the metabolome. For example,

we observed significantly higher concentrations of most LCFAs and fatty alcohols when comparing the vaginal metabolites of CT+ and CT+/MG+ to uninfected women both before and after adjustment for CST. This effect, however, appeared to be attenuated by MG co-infection: CT+/MG+ women had several LCFAs and fatty alcohols that were lower compared to CT+ women.

It is possible that the higher levels of LCFAs are related to CT survival, growth, and/or pathogenesis. Chlamydiae are obligate intracellular organisms that are auxotrophic for a variety of essential metabolites, including many lipids that they acquire from the host cell. Lipids are trafficked from various sources within the host cell into the inclusion and to the envelope of the reticulate bodies. The acquisition of lipids from host stores is thought to be critical for chlamydial replication and development³⁸. In eukaryotic cells, cholesterol esters, which are formed by the addition of a LCFA to the hydroxyl group of cholesterol, are the main transport and storage form of cholesterol³⁹. The enzyme acyl-CoA:cholesterol acyltransferase (ACAT) catalyzes the production of these esters⁴⁰. This process is important because in the un-esterified form, cholesterol can be toxic to the cell. A recent report⁴⁰ showed that chlamydial growth can be inhibited by blocking ACAT with a specific inhibitor, 4-hydroxycinnamic acid amide. Addition of a cholesterol ester (cholesterol esterified with a LCFA) rescued ACAT dependent inhibition⁴⁰.

These data raise the question of whether CT may be upregulating host production of LCFAs and cholesterol esters as part of a survival or growth mechanism. The relatively lower levels of lipid and lipid derivatives in CT/MG co-infected women versus uninfected women may reflect a direct or indirect competition between CT and MG for host lipids. Alternatively, the higher fatty alcohols and LCFAs may reflect greater wax ester production. Wax esters, a type of neutral lipid, are formed by the esterification of a fatty alcohol with a fatty acid derived from either acyl-CoA or acyl-ACP⁴¹. Neutral lipids are commonly found in vaginal fluid, along with secretions from the cervix, squamous epithelial cells and material from the sebaceous, apocrine and Bartholin's glands⁴². Thus, it is possible that the higher abundance of fatty acids, LCFAs and fatty alcohols reflects greater vaginal discharge associated with chlamydial infections, whether due to increased polymorphonuclear leukocytes or increased secretions from host cells. These questions are intriguing, but speculative. Larger, mechanistic studies will be required to further explore them.

No studies have examined the vaginal metabolome in the context of MG mono- or CT/MG co-infection, and only one other report³¹ has examined the vaginal metabolome in the context of CT infection. That study found fewer distinguishing metabolites but had a number of important differences from ours, including a smaller sample size, the use of 1H-NMR rather than GC-MS to detect a total of only 54 metabolites and a lack of testing for MG. Furthermore, the study design was quite different: CT infected women in the Parolin study were specifically selected not to have BV. Importantly, by controlling for CST rather than eliminating subjects with CST IV microbiota, our results may capture important changes that occur as a result of complex interactions between the host, CT and dysbiotic vaginal microbiota which would be missed by an approach that excludes CT mono-infected women with BV.

There are several limitations to our study. First, while we have the largest group of CT+ women studied to date in a metabolomic analysis, the CT+/MG+ co-infected group was small, limiting our ability to control for other potential confounders⁸. Secondly, these data are cross-sectional, with samples taken at the time of CT and CT/MG infection. Therefore, causal inferences (e.g. regarding the impact of the vaginal metabolome on susceptibility to CT or CT/MG co-infection) are not possible; nor is it possible to account for short term fluctuations in the vaginal microbiota. We do not have information on whether these represented first infections or re-infections. Cases were younger, less likely to use condoms, and reported more recent sexual partners than controls. While these factors were not found in univariate analysis to influence our outcome (i.e. the vaginal metabolome), we cannot completely exclude the possibility that these differences (or other unmeasured factors, such as *Candida* colonization) may impact the relationship between infection status and the metabolome. A large number of samples had to be excluded due to insufficient sample for metabolomics testing, unavailable 16S rRNA gene sequencing data or co-infection with *N. gonorrhoeae* or *Trichomonas vaginalis*, and there may have been unmeasured differences in these excluded samples. Batch effects could have introduced bias. Although we took a broad, untargeted GC-MS based approach, the library of identified metabolites does not reflect all the small metabolites present within the vaginal environment, and we may have missed crucial metabolites in our analysis. Importantly, the use of untargeted GC-based mass spectrometry excluded the detection of volatile fatty acids such as trimethylamine, acetate, or butyrate. The vaginal samples were stored in Amies transport medium. A number of metabolites, including lactic acid, were found in the transport medium and therefore had to be removed from analysis. Given limited sample size, we were not able to assess associations between relative abundance of specific bacterial species and vaginal metabolites. Finally, given the limitations of study design, we are only able to identify broad associations - it is not possible to untangle from this data what specific factors, or interaction between factors (i.e. host, vaginal microbiota and pathogen) are impacting metabolites and ultimately how this may affect susceptibility to, clearance of, or morbidity from CT and MG infections.

Despite these limitations, our study shows that a distinct vaginal metabolic profile distinguishes CT+ and CT+/MG+ from uninfected women, with significant overlap between CT+ and CT+/MG+ women. This study provides important foundational data documenting cross-sectional differences in the vaginal metabolome. Larger, mechanistic studies will be needed to explore the complex interactions between host, vaginal microbiota and pathogens, in order to better understand the pathogenesis of CT and CT/MG co-infection.

Methods

Study population and sample collection. Patient samples were drawn from three separate cohorts. CT positive cases (CT+) were drawn from the first (CT infected) visit of the Chlamydia Adolescent/Adult Reproductive Management (CHARM) study⁴³. CHARM recruited predominantly African-American CT+ young women of reproductive age in Baltimore, MD and followed them over time after treatment for CT. Women were

also tested for gonorrhea (GC) via NAAT and trichomoniasis via wet mount microscopy. Race-matched uninfected controls were drawn from enrollment visits of the Hormonal Contraception Longitudinal (HCL) study, (a cohort of low-risk women of reproductive age starting and stopping hormonal contraception recruited in Baltimore, MD⁴⁴) and the Human Microbiome Project Prospective Women's Health study at UMB (UMB-HMP), (a cohort of women of reproductive age recruited from clinics in Alabama and followed daily for 10 weeks)⁴⁵. Women in HCL and UMB-HMP were tested for CT, GC (via NAAT). Trichomoniasis was tested for via microscopy in HCL and inPouch in UMB-HMP. In UMB-HMP, if diagnosis of an STI or symptomatic condition was made at baseline, it was treated, and subjects were offered enrollment 30 days after treatment completion. Subjects were tested again for gonorrhea, chlamydia and trichomoniasis at their enrollment visit. All samples were self-collected mid-vaginal Copan Eswab specimens placed in Amies transport media and frozen at -80°C until use.

Ethics statement. All participants provided written informed consent. Ethical approval was obtained from the Institutional Review Boards of Johns Hopkins University, the University of Maryland Baltimore (UMB), and Montana State University, and all research was conducted in compliance with relevant guidelines and regulations.

Taxonomic assignment, community state type profiling and M. genitalium detection. Genomic DNA was extracted from 300 μl aliquots of re-suspended Amies solution as previously reported for the respective papers: (CT+ and CT+/MG+ cases⁴⁶), (controls recruited from Alabama⁴⁷), and (controls recruited from Maryland⁴⁴). For the QS DSP Midi kit extraction, samples were thawed on ice and a 500 μl aliquot from the vaginal swab was used as input⁴⁸. For the MagAttract kit, samples were thawed on ice and a 200 μl aliquot from the vaginal swab was used as input following the manufacturer protocol. Cells were lysed via shaking with 1.0 mm glass beads on the TissueLyser (Qiagen) at 20 Hz for 20 minutes and the final elution volume was 110 μl . Negative controls of water were extracted in the same manner as samples.

The vaginal microbiota was characterized by sequencing the V3-V4 regions of the 16S rRNA gene. Library construction was performed using either a 1-step protocol with sequencing carried out on the Illumina MiSeq platform for 600 cycles, resulting in 2×300 bp read lengths, or a 2-step PCR protocol, sequencing was carried out on the Illumina HiSeq. 2500 platform using Rapid Run Chemistry, and sequence data were processed according to Holm *et al.*⁴⁸. Amplicon sequence variants (ASVs) generated by DADA2 were classified taxonomically using the RDP Naïve Bayesian Classifier⁴⁹ trained with the SILVA v128 16S rRNA gene database^{48,50}. ASVs of major vaginal taxa were assigned species-level annotations using speciateIT (<http://ravel-lab.org/speciateit/>). Taxonomic data from a pool of over 12,000 vaginal samples were used to assign community state types (CSTs). Low-abundant taxa present at less than $10^{-5.5}$ abundance across the project were removed and samples with fewer than 5,000 reads were removed from analysis. Hierarchical clustering based on Jensen-Shannon distances between samples and Ward linkage was used to determine clusters for CST assignment. Four CSTs dominated by *Lactobacillus* spp. were identified: CST I (*L. crispatus*), CST II (*L. gasseri*), CST III (*L. iners*), and CST V (*L. jensenii*)⁵¹. CST IV contained more diverse organisms with low levels of *Lactobacillus* spp., (See Supplementary Fig. S1 for heatmap of 25 most abundant taxa with CST groupings).

Finally, DNA extracted from vaginal swabs was tested for MG using SpeeDx, a multiplex qPCR DNA based assay for detection of MG and resistance to macrolides by detecting one or more of 5 mutations in the 23S rRNA gene^{52,53}.

Sample selection and preparation for metabolomics. 289 samples were tested for MG: 157 CT+ cases from CHARM and 132 race matched CT- controls from HCL and HMP. Two control (i.e. CT-) samples tested positive for MG and were excluded. After exclusion of additional samples due to insufficient sample for metabolomic testing, unavailable 16S rRNA gene sequencing data, or co-infection with GC or *Trichomonas*, samples from 145 women were available for analysis (77 uninfected, 54 CT+ and 14 CT+/MG+ samples). N = 25 of the controls were recruited in Alabama, the remaining cases and controls were recruited from Maryland.

For quality control, an aliquot of 10 μl was taken from each sample and mixed in 10 replicates. Additionally, two samples of sterile Amies transport medium (controls) were run in duplicate. Samples, Amies controls, and QC supernatants were dried in a vacuum and derivatized with 50 μl methoxyamine hydrochloride (Sigma-Aldrich, MO, USA) (40 mg/mL⁻¹ in pyridine) for 60 min at 50°C , then with 50 μl MSTFA + 1%TMCS (Thermo, MA, USA) at 70°C for 120 min, followed by a 2-hour incubation at room temperature. 30 μl of 1 mg/ml hentriacontanoic acid was added to each sample prior to derivatization for use as an internal standard for normalization.

Metabolite profiles were acquired using a gas-chromatography mass-spectrometry (GC-MS) system (Agilent Inc, CA, USA) consisting of an Agilent 7890 gas chromatograph, an Agilent 5975 MSD and 7683B autosampler, as previously described⁵⁴. Briefly, gas chromatography was performed on a ZB-5MS (60 m \times 0.32 mm I.D. and 0.25 mm film thickness) capillary column (Phenomenex, CA, USA). The inlet and MS interface temperatures were 250°C , and the ion source temperature was adjusted to 230°C . An aliquot of 1 ml was injected with the split ratio of 10:1. The helium carrier gas was kept at a constant flow rate of 2.4 ml/min. The temperature program was: 5-min isothermal heating at 70°C , followed by an oven temperature increase of $5^{\circ}\text{C}/\text{min}$ to $^{\circ}\text{C}$ after which a final 10 min incubation at 310°C was performed. The mass spectrometer was operated in positive electron impact mode (EI) at 69.9 eV ionization energy at m/z 30–800 scan range. The spectra of all chromatogram peaks were evaluated using the AMDIS 2.71 (NIST, MD, USA) using a custom-built database (460 unique metabolites) of the University of Illinois Carver Metabolomics Center. Throughout the sample preparation, data-acquisition and data-preprocessing, samples were compared to the QCs to evaluate potential variation that may have arisen in the dataset throughout the analytical study. All known artificial peaks were identified and removed prior to data mining. To allow comparison between samples, all data were normalized to the internal standard in each chromatogram and sample volume. The instrument variability was within the standard acceptance limit (5%).

Metabolite pre-processing and statistical analyses. Metabolites detected in Amies (See Supplementary Table S5) were removed from analysis. Metabolites were excluded from analysis if they were undetected in $\geq 90\%$ of samples in all biological groups (CT+, CT+/MG+ and uninfected). Remaining missing values were assumed to be below the level of detection and were imputed (across all samples, irrespective of biological groups) using one half the minimum value obtained for a given metabolite^{55–58}. The final dataset available for analysis consisted of 163 metabolites. Following imputation of missing values, data were mean-centered, and a generalized log-transformation was applied to meet assumptions of linearity and normality required for regression models. Mean centering and generalized log-transformations were performed using MetaboAnalyst 4.0⁵⁸. Metabolic super and sub pathways were identified using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.

Unadjusted associations between infection status and demographic and behavioral factors were assessed using Fisher's exact test (Supplementary Table S1). Samples were visualized using a principal component analysis (PCA) constructed in MetaboAnalyst 4.0 (Fig. 1)⁵⁸. A Bray-Curtis dissimilarity matrix was constructed following normalization of the 163 metabolites, and a permutational multivariate analysis of variances (PERMANOVA) was performed to evaluate the association of potential sources of variation, including demographic and behavioral factors, with the vaginal metabolome. The Bray-Curtis dissimilarity matrix and PERMANOVA were conducted using the package *vegan* in R statistical Software.

We fit a linear regression model of mean-centered, generalized log-transformed metabolites on infection status, with and without controlling for CST. For each model, a significance level $\alpha = 0.05$ was used. To correct for multiple comparisons, p-values were adjusted using the Storey false discovery rate in the package *qvalue*⁵⁹. Linear regression and false discovery rate were conducted using R Statistical Software. Fold changes (FC) were calculated using $FC_i = \exp(\text{mean}(\log(X_i)) - \text{mean}(\log(Y_i)))$ where i is the metabolite of interest for the mean value of group of interest (X) and mean value of reference group (Y). Fold change values < 1 indicate lower metabolite concentration group X than in group Y (e.g. CT infected versus uninfected). Metabolite fold changes were calculated for (1) CT mono-infected vs. uninfected women, (2) CT/MG co-infected vs. uninfected women, and (3) CT/MG co-infected vs. CT mono-infected women. We additionally subset the data and calculated the metabolite fold change between infection states within CST III and CST IV. CSTs I, II, and V were not included as only two CT mono-infected, and no CT/MG co-infected women had vaginal microbiota corresponding to these CSTs, limiting our ability to make comparisons to uninfected women.

We performed partial least squares discriminant analysis (PLS-DA) on the mean centered, generalized log-transformed metabolite data. PLS-DA models were assessed based upon variance (R^2) and 10-fold cross-validation, which provides a quality assessment Q^2 with values close to 1 indicate reliable models and negative values indicate model overfitting. To assess whether differences between the groups were significant, a permutation test using the sum of squares between/sum of squares within (B/W) ratio was calculated. Variable importance in projection (VIP) score plots were used to display the top ten discriminatory metabolites. VIP scores ≥ 2.5 were considered important within a given model. PLS-DA, and VIP score plots were computed using MetaboAnalyst 4.0.

Data availability

The current 16S rRNA gene sequencing data for the cases can be found in SRA at: PRJNA509676. The 16S rRNA gene sequencing data for the controls recruited from Alabama can be found in SRA at: PRJNA208535. The 16S rRNA gene sequencing data for the controls recruited from Baltimore is currently in process and will be accessible through dbGaP. All data is available from the authors on reasonable request.

Received: 30 August 2019; Accepted: 3 February 2020;

Published online: 25 February 2020

References

- Molenaar, M. C., Singer, M. & Ouburg, S. The two-sided role of the vaginal microbiome in Chlamydia trachomatis and Mycoplasma genitalium pathogenesis. *J. Reprod. Immunol.* **130**, 11–17, <https://doi.org/10.1016/j.jri.2018.08.006> (2018).
- Aldunate, M. *et al.* Antimicrobial and immune modulatory effects of lactic acid and short chain fatty acids produced by vaginal microbiota associated with eubiosis and bacterial vaginosis. *Front. Physiol.* **6**, 164, <https://doi.org/10.3389/fphys.2015.00164> (2015).
- McKinnon, L. R. *et al.* The Evolving Facets of Bacterial Vaginosis: Implications for HIV Transmission. *AIDS Res. Hum. Retroviruses* **35**, 219–228, <https://doi.org/10.1089/AID.2018.0304> (2019).
- Boskey, E. R., Cone, R. A., Whaley, K. J. & Moench, T. R. Origins of vaginal acidity: high D/L lactate ratio is consistent with bacteria being the primary source. *Hum. Reprod.* **16**, 1809–1813 (2001).
- Gong, Z., Luna, Y., Yu, P. & Fan, H. Lactobacilli inactivate Chlamydia trachomatis through lactic acid but not H₂O₂. *PLoS One* **9**, e107758, <https://doi.org/10.1371/journal.pone.0107758> (2014).
- Tuddenham, S. & Ghanem, K. G. A microbiome variable in the HIV-prevention equation. *Sci.* **356**, 907–908, <https://doi.org/10.1126/science.aan6103> (2017).
- Martin, D. H. & Marrazzo, J. M. The Vaginal Microbiome: Current Understanding and Future Directions. *J. Infect. Dis.* **214**(Suppl 1), S36–41, <https://doi.org/10.1093/infdis/jiw184> (2016).
- Nelson, T. *et al.* Vaginal Biogenic Amines: Biomarkers of Bacterial Vaginosis or Precursors to Vaginal Dysbiosis? *Frontiers in Physiology* **6** (2015).
- Srinivasan, S. *et al.* Metabolic signatures of bacterial vaginosis. *MBio* **6**, <https://doi.org/10.1128/mBio.00204-15> (2015).
- Yeoman, C. J. *et al.* A multi-omic systems-based approach reveals metabolic markers of bacterial vaginosis and insight into the disease. *PLoS One* **8**, e56111, <https://doi.org/10.1371/journal.pone.0056111> (2013).
- Vitali, B. *et al.* Vaginal microbiome and metabolome highlight specific signatures of bacterial vaginosis. *Eur. J. Clin. Microbiol. Infect. Dis.* **34**, 2367–2376, <https://doi.org/10.1007/s10096-015-2490-y> (2015).
- Laghi, L. *et al.* Rifaximin modulates the vaginal microbiome and metabolome in women affected by bacterial vaginosis. *Antimicrob. Agents Chemother.* **58**, 3411–3420, <https://doi.org/10.1128/AAC.02469-14> (2014).

13. Ziklo, N., Huston, W. M., Taing, K., Katouli, M. & Timms, P. *In vitro* rescue of genital strains of Chlamydia trachomatis from interferon-gamma and tryptophan depletion with indole-positive, but not indole-negative Prevotella spp. *BMC Microbiol.* **16**, 286, <https://doi.org/10.1186/s12866-016-0903-4> (2016).
14. Ziklo, N., Vidgen, M. E., Taing, K., Huston, W. M. & Timms, P. Dysbiosis of the Vaginal Microbiota and Higher Vaginal Kynurenine/Tryptophan Ratio Reveals an Association with Chlamydia trachomatis Genital Infections. *Front. Cell Infect. Microbiol.* **8**, 1, <https://doi.org/10.3389/fcimb.2018.00001> (2018).
15. Jordan, S. J. *et al.* Lower Levels of Cervicovaginal Tryptophan Are Associated With Natural Clearance of Chlamydia in Women. *J. Infect. Dis.* **215**, 1888–1892, <https://doi.org/10.1093/infdis/jix240> (2017).
16. Borgogna, J. C. *et al.* The vaginal metabolome and microbiota of cervical HPV-positive and HPV-negative women: a cross-sectional analysis. *BJOG: An Int. J. Obstet. Gynaecol.* **127**, 182–192, <https://doi.org/10.1111/1471-0528.15981> (2020).
17. Onderdonk, A. B., Delaney, M. L. & Fichorova, R. N. The Human Microbiome during Bacterial Vaginosis. *Clin. microbiology Rev.* **29**, 223–238, <https://doi.org/10.1128/CMR.00075-15> (2016).
18. Lokken, E. M. *et al.* Association of Recent Bacterial Vaginosis With Acquisition of Mycoplasma genitalium. *Am. J. Epidemiol.* **186**, 194–201, <https://doi.org/10.1093/aje/kwx043> (2017).
19. Brotman, R. M. *et al.* Bacterial vaginosis assessed by gram stain and diminished colonization resistance to incident gonococcal, chlamydial, and trichomonal genital infection. *J. Infect. Dis.* **202**, 1907–1915 (2010).
20. Ness, R. B. *et al.* Bacterial vaginosis and risk of pelvic inflammatory disease. *Obstet. Gynecol.* **104**, 761–769, <https://doi.org/10.1097/01.AOG.0000139512.37582.17> (2004).
21. Ness, R. B. *et al.* A cluster analysis of bacterial vaginosis-associated microflora and pelvic inflammatory disease. *Am. J. Epidemiol.* **162**, 585–590, <https://doi.org/10.1093/aje/kwi243> (2005).
22. Torrone, E., Papp, J. & Weinstock, H. Prevalence of Chlamydia trachomatis Genital Infection Among Persons Aged 14–39 years — United States, 2007–2012. *Morbidity and Mortality Weekly Report* **2014**, 834–838 (2014).
23. Lis, R., Rowhani-Rahbar, A. & Manhart, L. E. Mycoplasma genitalium infection and female reproductive tract disease: a meta-analysis. *Clin. Infect. Dis.* **61**, 418–426, <https://doi.org/10.1093/cid/civ312> (2015).
24. Wiesenfeld, H. C. & Manhart, L. E. Mycoplasma genitalium in Women: Current Knowledge and Research Priorities for This Recently Emerged Pathogen. *J. Infect. Dis.* **216**, S389–S395, <https://doi.org/10.1093/infdis/jix198> (2017).
25. Taylor, B. D. *et al.* 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.* **94**, 414–420, <https://doi.org/10.1136/sextrans-2017-053376> (2018).
26. Workowski, K. A. Centers for Disease Control and Prevention Sexually Transmitted Diseases Treatment Guidelines. *Clin. Infect. Dis.* **61**(Suppl 8), S759–762, <https://doi.org/10.1093/cid/civ771> (2015).
27. Gaydos, C., Maldeis, N. E., Hardick, A., Hardick, J. & Quinn, T. C. Mycoplasma genitalium as a contributor to the multiple etiologies of cervicitis in women attending sexually transmitted disease clinics. *Sexually transmitted Dis.* **36**, 598–606, <https://doi.org/10.1097/OLQ.0b013e3181b01948> (2009).
28. Huppert, J. S. *et al.* Mycoplasma genitalium detected by transcription-mediated amplification is associated with Chlamydia trachomatis in adolescent women. *Sexually transmitted Dis.* **35**, 250–254, <https://doi.org/10.1097/OLQ.0b013e31815abac6> (2008).
29. Ravel, J. *et al.* Vaginal microbiome of reproductive-age women. *Proc. Natl Acad. Sci. USA* **108**(Suppl 1), 4680–4687, <https://doi.org/10.1073/pnas.1002611107> (2011).
30. Tamarelle, J. *et al.* The vaginal microbiota and its association with human papillomavirus, Chlamydia trachomatis, Neisseria gonorrhoeae and Mycoplasma genitalium infections: a systematic review and meta-analysis. *Clin. Microbiol. Infect.* **25**, 35–47, <https://doi.org/10.1016/j.cmi.2018.04.019> (2019).
31. Parolin, C. *et al.* Insights Into Vaginal Bacterial Communities and Metabolic Profiles of Chlamydia trachomatis Infection: Positioning Between Eubiosis and Dysbiosis. *Front. Microbiol.* **9**, 600, <https://doi.org/10.3389/fmicb.2018.00600> (2018).
32. Borgogna, J.-L. C. & Yeoman, C. J. In *The Human Microbiome Methods in Microbiology* 37–91 (2017).
33. Bautista, C. T. *et al.* Bacterial vaginosis: a synthesis of the literature on etiology, prevalence, risk factors, and relationship with chlamydia and gonorrhoea infections. *Mil. Med. Res.* **3**, 4, <https://doi.org/10.1186/s40779-016-0074-5> (2016).
34. Kanjee, U. & Houry, W. A. Mechanisms of acid resistance in Escherichia coli. *Annu. Rev. Microbiol.* **67**, 65–81, <https://doi.org/10.1146/annurev-micro-092412-155708> (2013).
35. Shah, P. & Swiatlo, E. A multifaceted role for polyamines in bacterial pathogens. *Mol. Microbiol.* **68**, 4–16, <https://doi.org/10.1111/j.1365-2958.2008.06126.x> (2008).
36. Gong, Z. *et al.* Arginine- and Polyamine-Induced Lactic Acid Resistance in Neisseria gonorrhoeae. *PLoS One* **11**, e0147637, <https://doi.org/10.1371/journal.pone.0147637> (2016).
37. Goytia, M. & Shafer, W. M. Polyamines can increase resistance of Neisseria gonorrhoeae to mediators of the innate human host defense. *Infect. Immun.* **78**, 3187–3195, <https://doi.org/10.1128/IAI.01301-09> (2010).
38. Cocchiario, J. L. & Valdivia, R. H. New insights into Chlamydia intracellular survival mechanisms. *Cell Microbiol.* **11**, 1571–1578, <https://doi.org/10.1111/j.1462-5822.2009.01364.x> (2009).
39. Cox, R. A. & Garcia-Palmieri, M. R. In *Clinical Methods: The History, Physical, and Laboratory Examinations* (eds. Walker, H. K., Hall, W. D. & Hurst, J. W.) (1990).
40. Peters, J. & Byrne, G. I. Chlamydia trachomatis growth depends on eukaryotic cholesterol esterification and is affected by Acyl-CoA:cholesterol acyltransferase inhibition. *Pathog. Dis.* **73**, ftv028, <https://doi.org/10.1093/femspd/ftv028> (2015).
41. Hamilton, R. J. R. *Waxes: chemistry, molecular biology and functions.* (Oily Press, 1995).
42. Huggins, G. R. & Preti, G. Volatile constituents of human vaginal secretions. *Am. J. Obstet. Gynecol.* **126**, 129–136 (1976).
43. Mark, K., *et al.* Chlamydia in Adolescent/Adult Reproductive Management Trial Study (CHARM): Clinical Core Protocol. *Contemporary Clinical Trials Communications* (2019).
44. Tuddenham, S. *et al.* Associations between dietary micronutrient intake and molecular-Bacterial Vaginosis. *Reprod. Health* **16**, 151, <https://doi.org/10.1186/s12978-019-0814-6> (2019).
45. Ravel, J. *et al.* Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. *Microbiome* **1**, 29, <https://doi.org/10.1186/2049-2618-1-29> (2013).
46. Edwards, V. L. *et al.* The Cervicovaginal Microbiota-Host Interaction Modulates Chlamydia trachomatis Infection. *MBio* **10**, <https://doi.org/10.1128/mBio.01548-19> (2019).
47. Ravel, J. *et al.* Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. *Microbiome* **1**, 29, <https://doi.org/10.1186/2049-2618-1-29> (2013).
48. Holm, J. B. *et al.* Ultra-high throughput multiplexing and sequencing of >500 bp amplicon regions on the Illumina HiSeq2500 platform (bioRxiv, 2018).
49. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
50. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–596, <https://doi.org/10.1093/nar/gks1219> (2013).
51. Ravel, J. In *Keystone Conference: Role of the Genital Tract Microbiome in Sexual and Reproductive Health.*
52. Tabrizi, S. N. *et al.* Prospective Evaluation of ResistancePlus MG, a New Multiplex Quantitative PCR Assay for Detection of Mycoplasma genitalium and Macrolide Resistance. *J. Clin. Microbiol.* **55**, 1915–1919, <https://doi.org/10.1128/JCM.02312-16> (2017).

53. Tabrizi, S. N. *et al.* Multiplex Assay for Simultaneous Detection of Mycoplasma genitalium and Macrolide Resistance Using PlexZyme and PlexPrime Technology. *PLoS One* **11**, e0156740, <https://doi.org/10.1371/journal.pone.0156740> (2016).
54. Ding, J. *et al.* Enhancement of gamma-aminobutyric acid (GABA) and other health-related metabolites in germinated red rice (*Oryza sativa* L.) by ultrasonication. *Ultrason. Sonochem.* **40**, 791–797, <https://doi.org/10.1016/j.ultsonch.2017.08.029> (2018).
55. Yang, J., Zhao, X., Lu, X., Lin, X. & Xu, G. A data preprocessing strategy for metabolomics to reduce the mask effect in data analysis. *Front. Mol. Biosci.* **2**, 4, <https://doi.org/10.3389/fmolb.2015.00004> (2015).
56. Smilde, A. K., van der Werf, M. J., Bijlsma, S., van der Werff-van der Vat, B. J. & Jellema, R. H. Fusion of mass spectrometry-based metabolomics data. *Anal. Chem.* **77**, 6729–6736, <https://doi.org/10.1021/ac051080y> (2005).
57. Wei, R. *et al.* Missing Value Imputation Approach for Mass Spectrometry-based Metabolomics Data. *Sci. Rep.* **8**, 663, <https://doi.org/10.1038/s41598-017-19120-0> (2018).
58. Xia, J. & Wishart, D. S. Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nat. Protoc.* **6**, 743–760, <https://doi.org/10.1038/nprot.2011.319> (2011).
59. Storey, J. D. A Direct Approach to False Discovery Rates. *J. R. Stat. Society. Ser. B* **64**, 479–498 (2002).

Acknowledgements

This research was supported by a developmental research grant awarded as a substudy of U19 AI08044, PI: Bavoil and Ravel. S.T. is supported by K23AI125715, PI: Tuddenham. Microbiome testing of HCL samples was supported by R01AI089878 (PI: Ghanem).

Author contributions

S.T., R.B. and K.G. conceived of the study. J.H. and C.G. conducted the MG testing. C.R. organized and collected the samples as well as the microbiome data, A.U. conducted the metabolomics testing. J.B., S.T., C.Y., M.S. and H.K. conducted the analysis. J.B., S.T. and C.Y. drafted the paper. All authors, including J.R., P.B., K.G. and R.B. contributed significantly to the methods and to revising the manuscript.

Competing interests

S.T. has received compensation as a consultant for Biofire Diagnostics and as a consultant and speaker for Roche Molecular Diagnostics. C.G. has received research funding and speaker fees from Hologic. C.J.Y. has received compensation as a consultant for Toltec Pharmaceuticals. R.M.B. received compensation from Evofem. J.R. is co-founder of LUCA Biologics, a biotechnology company focusing on translating microbiome research into live biotherapeutics drugs for women's health and owns stocks in the company. All other authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-60179-z>.

Correspondence and requests for materials should be addressed to S.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020