

# SCIENTIFIC REPORTS

OPEN

## Cigarette smoking is associated with an altered vaginal tract metabolomic profile

T. M. Nelson<sup>1,2</sup>, J. C. Borgogna<sup>1,2</sup>, R. D. Michalek<sup>3</sup>, D. W. Roberts<sup>4</sup>, J. M. Rath<sup>5,6</sup>, E. D. Glover<sup>5</sup>, J. Ravel<sup>7,8</sup>, M. D. Shardell<sup>9</sup>, C. J. Yeoman<sup>1,2</sup> & R. M. Brotman<sup>7,10</sup>

Cigarette smoking has been associated with both the diagnosis of bacterial vaginosis (BV) and a vaginal microbiota lacking protective *Lactobacillus* spp. As the mechanism linking smoking with vaginal microbiota and BV is unclear, we sought to compare the vaginal metabolomes of smokers and non-smokers (17 smokers/19 non-smokers). Metabolomic profiles were determined by gas and liquid chromatography mass spectrometry in a cross-sectional study. Analysis of the 16S rRNA gene populations revealed samples clustered into three community state types (CSTs) --- CST-I (*L. crispatus*-dominated), CST-III (*L. iners*-dominated) or CST-IV (low-*Lactobacillus*). We identified 607 metabolites, including 12 that differed significantly ( $q$ -value  $< 0.05$ ) between smokers and non-smokers. Nicotine, and the breakdown metabolites cotinine and hydroxycotinine were substantially higher in smokers, as expected. Among women categorized to CST-IV, biogenic amines, including agmatine, cadaverine, putrescine, tryptamine and tyramine were substantially higher in smokers, while dipeptides were lower in smokers. These biogenic amines are known to affect the virulence of infective pathogens and contribute to vaginal malodor. Our data suggest that cigarette smoking is associated with differences in important vaginal metabolites, and women who smoke, and particularly women who are also depauperate for *Lactobacillus* spp., may have increased susceptibilities to urogenital infections and increased malodor.

According to the U.S. National Health Interview Survey, 13.9 percent of U.S. women reported that they smoked cigarettes in 2016<sup>1</sup>, and it has been well-documented that women who smoke have greater risks for adverse reproductive health outcomes, including premature birth, delivery of low birth weight infants, certain birth defects, and ectopic pregnancy<sup>2-6</sup>. Smokers are also more susceptible to bacterial infections than non-smokers<sup>7</sup>, however, few studies have investigated the mechanisms linking smoking and gynecologic infections. Smoking has been shown in large observational studies to be a dose-dependent risk factor for the diagnosis of bacterial vaginosis (BV)<sup>8-13</sup>, as well as significantly associated with risk of other genital infections in females including *Trichomonas vaginalis*<sup>14</sup>, herpes simplex virus type 2 (HSV-2)<sup>15</sup> and *Chlamydia trachomatis*<sup>16,17</sup>. Smoking has long been implicated in higher risk for oral and genital human papillomavirus (HPV) prevalence and viral load, progression to cervical pre-cancer, and vulval intraepithelial neoplasia<sup>18-23</sup>. Smoking is associated with damaged cervical epithelium through DNA modification and suppressed local and systemic immune responses<sup>24</sup>, both of which may increase susceptibility to a wide range of female reproductive tract infections<sup>25,26</sup>. Nicotine's major metabolite, cotinine, has been shown to become concentrated in cervical mucus, providing evidence that smoking can directly affect the vaginal and cervical epithelium<sup>27,28</sup>. Smoking cessation was also associated with dramatic changes in the gut microbiota in one study<sup>29</sup>.

<sup>1</sup>Department of Animal and Range Sciences, Montana State University, Bozeman, MT, USA. <sup>2</sup>Department of Microbiology and Immunology, Montana State University, Bozeman, MT, USA. <sup>3</sup>Metabolon Inc, Durham, NC, USA. <sup>4</sup>Department of Ecology, Montana State University, Bozeman, MT, USA. <sup>5</sup>Department of Behavioral and Community Health, University of Maryland School of Public Health, College Park, MD, USA. <sup>6</sup>Truth Initiative, Washington DC, USA. <sup>7</sup>Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD, USA. <sup>8</sup>Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, MD, USA. <sup>9</sup>Translational Gerontology Branch, National Institute on Aging, Baltimore, MD, USA. <sup>10</sup>Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, MD, USA. Correspondence and requests for materials should be addressed to C.J.Y. (email: [carl.yeoman@montana.edu](mailto:carl.yeoman@montana.edu)) or R.M.B. (email: [rbrotman@som.umaryland.edu](mailto:rbrotman@som.umaryland.edu))

Variable	SS (trace)	Pseudo-F	p-value	Proportion (%)	Cumulative Proportion (%)
<b>Individual model test</b>					
CST	3181.00	9.94	0.002	22.62	NA
Smoking status	880.65	2.27	0.074	6.26	NA
Race	1120.00	2.94	0.033	7.96	NA
Education level	1216.90	3.22	0.038	8.65	NA
<b>Fitted model with selected variables, adjusted R<sup>2</sup> = 0.26</b>					
CST	3181.00	9.94	0.001	22.62	22.62
Race	817.71	2.68	0.037	5.82	28.44
Education level	533.18	1.79	0.11	3.79	32.23
Smoking status	292.20	0.98	0.37	2.08	34.31

**Table 1.** Significantly fitted predictor variables to vaginal metabolites. Distance based linear modelling (DISTLM) was performed on vaginal metabolites fitted with participant behavioral variables listed in Table 2. DISTLM was conducted with adjusted R<sup>2</sup> over 9,999 permutations. DISTLM identifies the best-fit model based on all the available variables that best-explain the composition of vaginal metabolites. The visual display of these data as represented by the distance-based redundancy analysis (dbrDA) plot is displayed in Fig. 3.

In a 2014 study, our research group confirmed that the composition of the vaginal microbiota is strongly associated with smoking<sup>30</sup>. In that study, we compared the vaginal microbiota, as determined by 16S rRNA gene amplicon sequencing, between smokers and non-smokers. We reported that women who had a vaginal microbiota that were lacking significant numbers of *Lactobacillus* spp. were 25-fold more likely to report current smoking than women with a *L. crispatus*-dominated microbiota. Most *Lactobacillus* spp. are thought to provide broad-spectrum protection to pathogenic infections through the production of lactic acid and bacteriocins<sup>31,32</sup>. The lactic acid reduces the vaginal pH to  $\leq 4.0$  and is a potent bactericide and virucide<sup>31,33–36</sup>. Indeed, epidemiologic studies have shown that a relatively low level of *Lactobacillus* spp. and the presence of a wide array of strict and facultative anaerobes in the vaginal microbiota, as observed in the clinical diagnosis of BV, is associated with increased incidence of HIV and other sexually transmitted infections (STIs)<sup>15,37–43</sup>.

Despite advances in understanding the epidemiologic links between smoking and women's reproductive health, the mechanistic processes by which smoking affects the vaginal microenvironment and BV is unclear. The effect of smoking on the function of the microbiome can be investigated by assessing the metabolome. The metabolome is the set of small (<1 kDa) molecule chemicals, which includes host and microbially-produced and modified biomolecules, as well as exogenous chemicals<sup>44</sup>. Measurement of the metabolomic profile from any biological sample is expected to contain numerous low molecular mass molecules with different physicochemical characteristics and concentration ranges<sup>44</sup>.

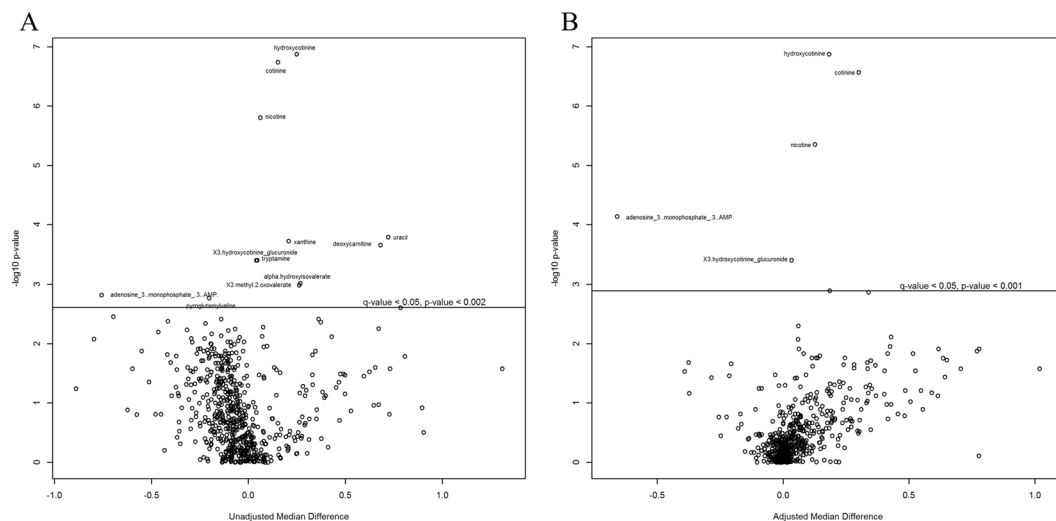
The metabolome is an important characteristic of the vaginal microenvironment and differences in some metabolites are associated with functional variations of the vaginal microbiota<sup>45–48</sup>. For example, the presence of certain chemicals or metabolites have been shown *in vitro* to directly reduce or inhibit the growth of select bacterial species<sup>49–51</sup>. Biosynthesis of biogenic amines (BAs) (cadaverine, putrescine, spermine, spermidine, trimethylamine, and tyramine) may allow various bacteria to survive in low pH environments<sup>52</sup>, like that found in the healthy *Lactobacillus*-dominated vagina. BAs have also been shown to enhance growth rates of various pathogenic bacteria, such as *Neisseria gonorrhoeae*, and shield them from host innate immunological defenses<sup>53–55</sup>. The recent applications of metabolomics with *in vivo* samples have identified relationships between bacterial species and vaginal metabolomic profiles. In particular *Dialister*, *Mobiluncus*, *Atopobium*, *Prevotella*, *Mycoplasma* and *Gardnerella* species were correlated with the presence of several metabolites along with vaginal odor and discharge<sup>45–48</sup>. We have hypothesized that BAs may be an important feature in the destabilization of *Lactobacillus* spp.-dominant vaginal microbiota and the initiation of BV, as well as the characteristic malodor of BV<sup>56</sup>. The decarboxylation of amino acid precursors to form BAs results in increased pH and may increase the risk for BV<sup>57</sup>.

In this study, we sought to characterize vaginal microbiota functional differences between smokers and non-smokers by investigating the vaginal metabolome. These data may allow enhanced understanding of the mechanism by which smoking may increase the risk of urogenital infections and may contribute to our understanding of the effect of smoking on the female reproductive tract.

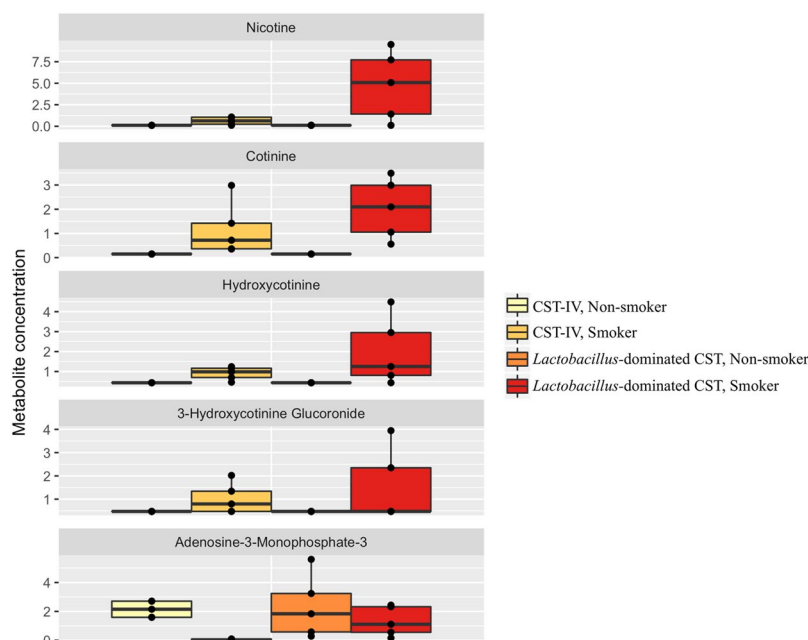
## Results

In total, 607 compounds were identified in the vaginal metabolome of 36 women. PCA indicated that the metabolomic profiles clustered by both smoking status and bacterial community state type (CST) (Figure S1). CST accounted for 23% of the metabolite variation observed in the complete data set (distance based linear model, DISTLM F = 9.9397, P<sub>PERM</sub> = 0.0001; Table 1), while smoking status explained 6%. However, the concentrations of 12 metabolites were significantly different between smokers and non-smokers (q-value = <0.05; Fig. 1A, Table S1). Tobacco constituents and their primary breakdown products (nicotine, cotinine and hydroxycotinine) represented the strongest differences between smokers and non-smokers with non-smokers having a 2- to 12-fold reduction in these compounds (q-value = <0.05). After adjusting for the influence of CST, five of the metabolites persisted in their differences between smokers and non-smokers (Figs 1B, 2).

Further, 142 metabolites had abundances that were marginally significantly different between smokers and non-smokers (q-value between 0.05 and 0.10) without adjustment for CST (Table S2). These compounds are



**Figure 1.** Compounds differing between smokers and non-smokers with and without adjustment for CST. Volcano plots display  $-\log_{10}$  (p-value) and the median difference in concentration between smokers and non-smokers when unadjusted (A) and adjusted for community state type (B). Quantile regression was conducted on centered and scaled metabolite concentrations. Significance testing was conducted with Wilcoxon rank sum test and corrected for multiple comparisons. Metabolites that differed significantly where  $q\text{-value} \leq 0.05$  between smokers and non-smokers are shown above the line in each plot.



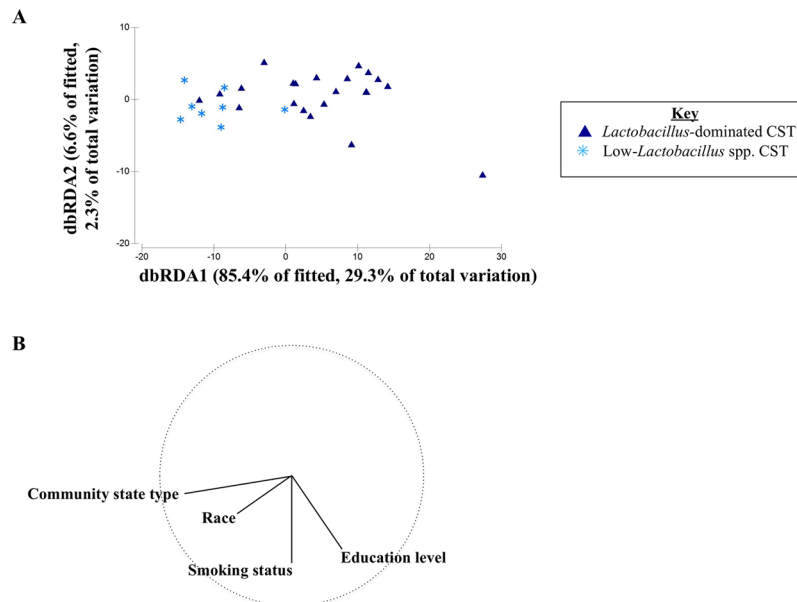
**Figure 2.** Vaginal metabolites that differ between smokers and non-smokers. Boxplots display metabolites identified as significantly ( $q\text{-value} < 0.05$ ) different in the vagina of smokers and non-smokers when unadjusted and adjusted for the impact of bacterial community state type (CST). Samples categorized as CST-I (*L. crispatus*-dominated) and CST-III (*L. iners*-dominated) were grouped and compared with CST-IV (low-*Lactobacillus* spp.). Quantile regression was conducted on centered and scaled metabolite concentrations. Significance testing was conducted with Wilcoxon rank sum test and corrected for multiple comparisons.

involved in amino acid metabolism, including some essential for bodily functions such as lysine, tryptophan, leucine, isoleucine, histidine and methionine as well as many amino acid dipeptides required for protein hydrolysis. Interestingly, the greatest differences in the composition and abundance of the metabolomes between smokers and non-smokers were among biogenic amines, including, cadaverine (fold change in non-smokers,  $FC = -53.19$ ), putrescine ( $FC = -15.80$ ), agmatine ( $FC = -15.49$ ), tyramine ( $FC = -5.15$ ) and tryptamine ( $FC = -4.43$ ), which were much higher in smokers (Table S2, Figure S2).

	Non-smoker		Smoker		p-value <sup>1</sup>
	n	%	n	%	
<b>Participant details</b>					
Age					0.011
19–28	16	44	4	11	
29–38	1	3	6	17	
39–48	2	6	7	19	
Marital status					0.066
Single, never married	17	47	8	22	
Separated, divorce, widowed	1	3	5	14	
Married	1	3	3	8	
Race					0.170
Asian/Pacific Islander	3	8	1	3	
White	8	22	4	11	
African American/black	4	11	10	28	
Hispanic	3	8	0	0	
Multi-racial	1	3	1	3	
Other	0	0	1	3	
Education level					1.000
High School, up to 12 years	0	0	1	3	
College and graduate, > 12 years	19	53	16	44	
<b>Clinical</b>					
CST					0.011
I, <i>L. crispatus</i> -dominated	13	36	4	11	
III, <i>L. iners</i> -dominated	4	11	3	8	
IV, Low- <i>Lactobacillus</i> spp.	2	6	10	28	
Nugent's Gram stain score					0.000
0–3	17	47	8	22	
4–6	2	6	2	6	
7–9	0	0	7	19	
Vaginal pH					0.004
<=4.0	10	28	4	11	
4.1–5.5	1	5	1	3	
4.6–5.0	4	11	2	6	
>=5.1	4	11	10	28	
<b>Self-reported symptoms</b>					
Vaginal odor					0.969
No	15	42	16	44	
Yes	4	11	1	3	
Vaginal irritation, 24 hours prior					1.000
No	19	53	17	47	
Yes	0	0	0	0	
Vaginal itching, 24 hours prior					1.000
No	19	53	17	47	
Yes	0	0	0	0	
Vaginal burning, 24 hours prior					1.000
No	19	53	17	47	
Yes	0	0	0	0	
Pain urinating, 24 hours prior					1.000
No	19	53	17	47	
Yes	0	0	0	0	
Vaginal discharge, 24 hours prior					0.847
No	14	39	14	39	
Yes	5	14	3	8	
Continued					

	Non-smoker		Smoker		p-value <sup>1</sup>
	n	%	n	%	
<b>Participant details</b>					
<b>Hygiene</b>					
Vaginal douche, 2 months prior					0.108
No	18	50	9	25	
Yes	1	3	7	19	
Not recorded	0	0	1	3	
Menstruating currently					0.075
No	9	25	9	25	
Yes	10	28	8	22	
Tampon or pad, 24 hours prior					0.456
No pad, no tampon	10	28	10	28	
Pad only	2	6	1	3	
Tampon only	4	11	2	6	
Tampon and pad	2	6	1	3	
Not recorded	1	3	3	8	
<b>Sexual behaviors</b>					
Lifetime number of sexual partners					0.25
0–6	13	36	3	8	
7+	6	17	12	33	
Not recorded	0	0	2	6	
Number of sex partners, 2 months prior					0.000
0	4	11	2	6	
1	12	33	14	39	
2	3	8	1	3	
Vaginal intercourse with a condom, 24 hours prior					0.673
No vaginal intercourse	14	39	9	25	
Vaginal intercourse no condom	3	8	6	17	
Vaginal intercourse with condom	2	6	2	6	
Anal intercourse, 24 hours prior					1.000
No	19	53	17	47	
Yes	0	0	0	0	
Sex toy use, 24 hours prior					1.000
No	19	53	17	47	
Yes	0	0	0	0	
Lubricant use, 24 hours prior					1.000
No	19	53	17	47	
Yes	0	0	0	0	
Partner type					0.000
Regular	10	28	12	33	
Occasional	2	6	0	0	
New	1	3	0	0	
Not recorded	6	17	5	14	
Receptive oral sex, 24 hours prior					0.543
Yes	2	6	1	3	
No	17	47	16	44	
Digital penetration, 24 hours prior					0.555
Yes	3	8	2	6	
No	16	44	15	42	
Thong use, 24 hours prior					0.000
Yes	6	17	3	8	
No	13	36	13	36	
Not recorded	0	0	1	3	

**Table 2.** Factors associated with smoking status, Baltimore, MD (n=36) <sup>1</sup>p-value determined using Fisher's exact test.



**Figure 3.** Participant predictor variables fitted to vaginal metabolites. Distance-based linear modelling (DISTLM) was conducted on log-transformed Euclidean distance matrix of vaginal metabolites. All variables were included in the final model that when applied to the data cloud of vaginal metabolites (Table 1). Distance-based redundancy analysis (dbRDA) plot (A) displays the metabolite composition fitted to the four variables labeled with the strongest variable, CST. Variable vectors indicating strength and direction are identified (B).

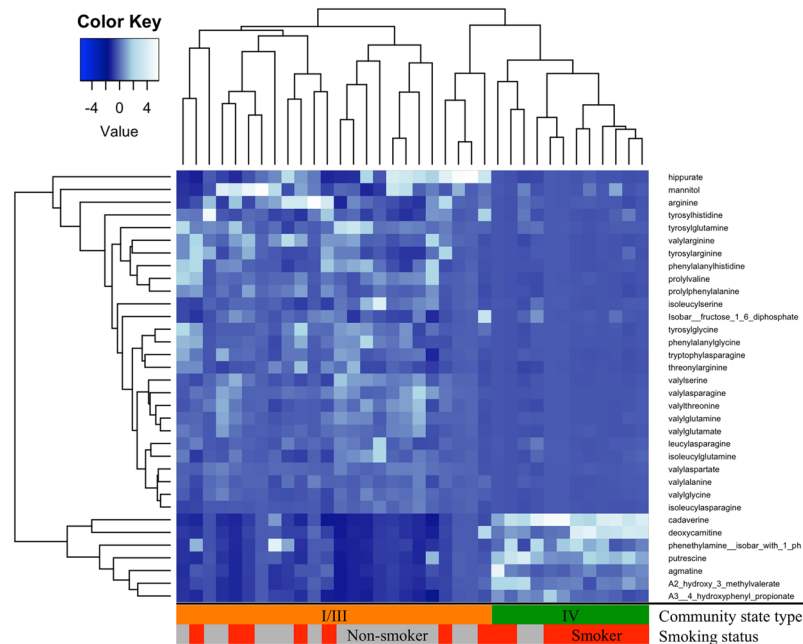
The associations observed between smoking and the vaginal metabolome may be explained in part by participant-specific variables, such as sexual behaviors, health, race and other confounding factors. We utilized a distance-based linear modeling (DISTLM) approach to explain the contribution of confounding participant variables to the vaginal metabolome (Table 1). Due to the relatively small sample size of our study, the model was limited to a total of four variables, which we selected as the best fit from all potential models with the inclusion of smoking status and CST. The model including CST, race, education level, and smoking status, had an adjusted  $R^2$  of 0.26, and these variables explained 34% of the observed variation in the vaginal metabolomic profiles (Table 1). Using this analytical approach based on the full metabolome, smoking status did not show a statistically significant relationship to the full spectrum of vaginal metabolites after adjustment for CST, race, and education, although in some cases, when the associations are large and the p-values do not reach significance, it may be due to the lack of power associated with the relatively small sample size (Fig. 3, Table 1).

In addition to the multivariate analyses above, we also sought to directly contrast the metabolome of *Lactobacillus*-dominated CST-I/III and the low-*Lactobacillus*/diverse anaerobes observed in CST-IV. To increase the power of significance testing, we combined *L. crispatus*-dominated CST-I and *L. iners*-dominated CST-III. In bivariate analysis, there was a significant difference ( $q$ -value  $\leq 0.05$ ) in the concentrations of 67% of all identified vaginal metabolites between CST-I/III and CST-IV (Table S3). After adjusting for the influence of a woman being a current smoker, the abundance of 54% of vaginal metabolites still differed significantly ( $q$ -value  $\leq 0.05$ ). The majority of the metabolites that were elevated in the *Lactobacillus*-dominated CSTs were amino acids, lipids, peptides and especially dipeptides (Fig. 4). Carbohydrates, mannitol, lactate, xylulose and fucose were higher in the *Lactobacillus*-dominated CSTs compared to CST-IV (Table S3). The glycine conjugate of benzoic acid, hippurate, was also markedly higher (34-fold) in *Lactobacillus*-dominated communities. The BAs (cadaverine, putrescine, agmatine, tyramine, tryptamine) and the straight chain fatty acid deoxycarnitine were also observed to be significantly higher in CST-IV compared to *Lactobacillus*-dominant communities (Table S3). Correlations with pH measurement and BA concentrations were positive, yet lacked statistical significance (Table S4, Figure S3).

Further, among women in CST-IV, the BAs were markedly higher in smokers versus non-smokers ( $q$ -value:  $< 0.05$ , Cadaverine (89-fold), Putrescine (26-fold), Agamatine (26-fold), Tyramine (10-fold) and Tryptamine (7-fold), Figure S2).

## Discussion

Our group and others have previously reported that female smokers are more likely to display a low-*Lactobacillus* CST-IV vaginal microbiota and are at increased risk for morbidities such as bacterial vaginosis<sup>9,30</sup> and sexually transmitted infections<sup>39</sup>. To obtain greater insight into this relationship, we assessed the vaginal metabolome of smokers and non-smokers. Using a comprehensive metabolomics approach and multiple analyses of the data, we identified an extensive and diverse range of vaginal metabolites for which profiles were affected by both the microbiology and smoking status. Bacterial composition (CST) was the most pronounced driver of the vaginal metabolome in our model and was associated with changes in 57% of all metabolites, suggesting vaginal microbiota are the major drivers of the vaginal metabolome. We therefore carefully controlled for CST in multivariate



**Figure 4.** Metabolites differ between CST. Heatmap displays selection of metabolites with a high fold change identified as significantly ( $q$ -value =  $<0.05$ ) different in the vagina of bacterial community state type (CST) when unadjusted and adjusted for the impact of smoking status. Low-*Lactobacillus* CST-IV and *Lactobacillus*-dominated CST-I and CST-III are grouped. Quantile regression was conducted on centered and scaled metabolite concentrations. Significance testing was conducted with Wilcoxon rank sum test and corrected for multiple comparisons.

analysis and also conducted stratified analysis based on CST so that we could directly contrast smokers versus non-smokers, while essentially holding CST constant. As expected, nicotine and its breakdown products were markedly elevated in the vagina of smokers. In humans, 70–80% of nicotine is converted to cotinine followed by conversion to hydroxycotinine<sup>58</sup>. Hydroxycotinine is a well-known biomarker identified in the urine, plasma and serum of individuals exposed to both active and passive smoking<sup>59–61</sup>. Previous studies have identified nicotine and cotinine in the cervical mucus of smokers<sup>27</sup> with good correlation to concentrations in blood and urine<sup>62,63</sup>. Our findings are consistent with the vaginal metabolome contributing to the mechanism linking smoking to the microbiome.

Another key finding was that we observed significant increases in the abundance of various BAs among smokers relative to non-smokers, which was far more pronounced in women with a low-*Lactobacillus* CST-IV. BAs are unique molecules that carry one or more amine groups ( $\text{NH}_2$ ). They are essential to mammalian and bacterial physiology, tightly controlled in cellular metabolisms. Cadaverine, putrescine, agmatine and tryptamine have roles in the metabolism of essential amino acids, including tryptophan and lysine. Their accumulation indicates upheaval or alterations in these metabolic systems. In particular, the odors of the amines cadaverine and putrescine are foul-smelling to humans, identifiable as indicators of tissue decomposition associated with death or bacterial contamination<sup>64,65</sup>. Several of these BAs, including cadaverine and putrescine, have previously been correlated with diagnosis of BV<sup>45,66,67</sup> and implicated in the associated ‘fishy’ vaginal malodor of BV<sup>46,47</sup>.

Our group recently suggested a hypothetical model for the displacement of vaginal *Lactobacillus* spp. and increased risk of BV and urogenital infection with BAs<sup>68</sup>. We hypothesize that in the vaginal canal, BAs may favor non-*Lactobacillus* species, while also increasing the vaginal pH, collectively enabling colonization by a more diverse community, as is observed with CST-IV. Our hypothesis is based on two observations. First, the amino-acid decarboxylation reactions that produce BAs involve the consumption of intracellular hydrogen ions and is a well-described bacterial acid resistance and mitigation mechanism<sup>52</sup>. The consumption of hydrogen ions increases the pH of the local habitat<sup>52</sup>, overcoming what is widely considered the primary barrier to pathogen outgrowth and a clinically-recognized symptom of BV. Second, the growth of several pathogens, including the urogenital pathogen, *Neisseria gonorrhoeae*, and their resistance to host immunological defenses has been shown to be superior in the presence of various BAs<sup>53–55,69</sup>. In this study, we observed that lactate was lower and BAs were higher, in low-*Lactobacillus* CST-IV participants as expected. However, lactate was higher in *Lactobacillus*-dominated CST-I/III even when BA concentrations were low. Lactate concentration includes both protonated lactic acid (LAH) and lactate anion ( $\text{LA}^-$ ) with the former being recognized as the active microbicidal form capable of inactivating BV-associated bacteria<sup>31,70</sup>. Lactate concentrations increase with increasing hydrogen ion concentrations<sup>31</sup> and therefore the reduction in lactate and increase in BAs further supports our hypotheses, and suggest that bacteria present in low-*Lactobacillus* CST-IV use available hydrogen ions to perform amino acid decarboxylation and produce BAs, thus resisting acid stress<sup>56</sup>.



Our results identifying the presence of nicotine breakdown products in smokers may reflect differences in the degree of transport and/or accumulation of nicotine and its derivatives in the vagina by CST. These differences may relate to the co-variation in pH with CST<sup>33</sup>. Absorption of nicotine across biological membranes (such as cervicovaginal endothelial/epithelial cells) has previously been shown to be pH dependent<sup>56,71</sup>. Nicotine is a weak base ( $pK_a = 8.0$ ) and in acidic environments nicotine does not rapidly cross membranes<sup>68</sup>. Cotinine and hydroxycotinine are more acidic with a  $pK_a = 4.8$  and  $4.3$ , respectively<sup>72</sup>. Previous studies which have measured both in serum and vaginal samples identified higher levels of nicotine in the cervical mucus compared with serum samples, yet cotinine values were similar<sup>27,63</sup>. Nicotine may be selectively concentrated in the vagina because the majority (72%) of women display a vaginal pH of 4.0 to 4.6<sup>33</sup>. Conversely, women with low-*Lactobacillus* CST-IV vaginal microbiota have a vaginal pH of  $5.3 \pm 0.6$ <sup>33</sup> and, consistently displayed comparatively lower levels of nicotine. Alternatively, these findings may indicate microbiological metabolism of nicotine in the more diverse and low-*Lactobacillus* state of CST-IV.

Similarly, hippurate, a normal excretory product of urine that is increased with exposure to phenolic compounds and toluene, a byproduct of cigarette smoke<sup>73</sup>, was increased in *Lactobacillus*-dominated CST-I/III over CST-IV and also in non-smokers over smokers. Hippurate is a known substrate of *Gardnerella vaginalis*<sup>74</sup>, an organism often found in higher abundances when *Lactobacillus* spp. are low and commonly associated with symptomatic BV<sup>75–77</sup>. Hippurate is decreased in the low-*Lactobacillus* CST-IV and this may suggest that microbial utilization of hippurate and therefore smoking may move the vaginal environment closer to favor the proliferation of *G. vaginalis*.

Beyond metabolites directly affiliated with nicotine metabolism, we observed substantive shifts in dipeptides and biogenic amines. More than 150 dipeptides in smokers or CST-IV participants were significantly decreased relative to non-smokers and those with *Lactobacillus* spp.-dominated vaginal microbiota. In a previous study, Ghartey and colleagues noted significant reductions in vaginal dipeptides among women who delivered infants preterm, for which CST-IV and BV are important risk factors<sup>78</sup>. Dipeptides refer to one or more amino acid joined by a peptide bond with important roles in protein metabolism and cell signaling. CST-I and III women had higher concentrations of dipeptides relative to CST-IV, suggesting *Lactobacillus* spp. dominance may be important to this phenotype. Dipeptides are constituents of the peptidoglycan cell wall of bacteria and are made during its' synthesis. Therefore, some dipeptides, such as muramyl dipeptides, serve as signal to the immune system in mammals and play a direct role in the regulation of inflammation<sup>79</sup>. Their production has also been indicated in *Lactobacillus* spp. and other bacteria as a mean of quorum sensing and cellular signaling<sup>80</sup>. *Lactobacillus* spp. may exploit dipeptides to increase their osmotolerance<sup>81</sup>, or produce dipeptides for amensalistic purposes. Various *Lactobacillus* spp. have been noted to produce cyclic dipeptides with anti-fungal<sup>80,82–84</sup>, anti-viral<sup>85</sup> and anti-bacterial<sup>80,86</sup> properties. This includes the production of dipeptides by a vaginal isolate of *L. reuteri* that disrupts the virulence capabilities of *Staphylococcus aureus* involved in toxic shock syndrome<sup>67</sup>. The increase in dipeptides in non-smokers and *Lactobacillus* spp.-dominant CSTs may be reflective of increased production by *Lactobacillus* spp. of these bioactive compounds. Conversely, the reduction in dipeptides in women with a low-*Lactobacillus* CST-IV vaginal microbiota may be a result of the increased proteolytic activity from bacteria present in this CST. Many BV-associated bacteria have been correlated with, or shown to be capable of secreting proteases, which break down proteins into amino acids<sup>87–91</sup>. Aside from reducing the number of detected dipeptides in the vaginal tract, proteases may also have the effect of inactivating proteins important to host defenses and make host tissues more susceptible to other organisms' virulence factors<sup>87</sup>.

The breakdown compounds of a number of drugs, such as cocaine (norbenzoylcegonine, benzoylcegonine), antidepressants (escitalopram), common painkillers (acetaminophen glucuronide, ibuprofen) and decongestants (pseudoephedrine) were each observed in one or more samples in the study. This indicates that some drugs could possibly be assessed from vaginal metabolomic profiles. There is relatively little known about the relationship between individual drugs and their impact on the vaginal microbiome. Cocaine use has been associated with shifts in bacterial phyla in the gut<sup>92</sup> and additionally has been associated with a greater likelihood of contracting sexually transmitted infections<sup>93,94</sup>. The use of antidepressants has been linked to menstrual disorders and hormonal changes in women<sup>95,96</sup> both of which may cause indirect shifts in the vaginal microbiota. As these studies may suggest a potential impact of drugs and medications and their byproducts on the vaginal microbiota, further exploration is needed to make any conclusions.

Our data are consistent with recent studies of the vaginal metabolome<sup>45,47,97</sup>. In prior studies, the BAs (cadaverine, putrescine and tyramine) were consistently higher in women with BV<sup>45,47,48,97</sup>. We did not detect trimethylamine in our samples as has previously been identified and we suspect this is due to its high volatility<sup>45,47,48,97</sup>. Srinivasan *et al.* reported the same issue with their lack of trimethylamine detection using similar methods yet reported lower levels of trimethylamine oxide (an intermediary product) in women with BV using separate methods<sup>97</sup>. Therefore, there are limitations to the use of metabolomics, including particle resolution, compound sensitivity, polarity and volatility<sup>98</sup>. Identification of detected metabolites is further constrained by comparison to facility-built databases<sup>98,99</sup>. However, this is best achieved by utilizing standardized collection, preparation and<sup>98</sup> pairing liquid chromatography (LC), gas chromatography (GC) and mass spectrometry (MS) which can enhance identification of analytes with differing characteristics<sup>98</sup>.

Aside from biogenic amines, McMillian *et al.* reported alpha-hydroxyisovalerate and gamma-hydroxybutyrate (GHB) as associated with BV and high bacterial diversity<sup>47</sup>. We also observed higher levels of alpha-hydroxyisovalerate and GHB in low-*Lactobacillus* CST-IV and additionally we observed a trend towards increases in smokers. Alpha-hydroxyisovalerate was positively correlated with diverse BV-associated bacteria such as *Atopobium*, *G. vaginalis*, *Dialister* and *Gemella*. McMillian *et al.* went on to show how *G. vaginalis*, a vaginal bacterial species commonly associated with BV is a producer of GHB<sup>47</sup>. Srinivasan *et al.* also identified 12-hydroxyicosatetraenoic acid (12-HETE) in cases of BV and we also detected a significantly high abundance of this compound in women from CST-IV over the *Lactobacillus*-dominant CSTs (I or III), although it was not associated with smoking status<sup>97</sup>.



A limitation of our study was the small sample size and the distribution of smokers and non-smokers within each community state type (17 CST-I, 7 CST-III, 12 CST-IV participants). As a result, we have relatively reduced power in some of these statistical tests in stratified analyses. Some of the assumptions of traditional parametric statistical tests are sensitive to small sample sizes, and the use of non-parametric statistical tests, as was employed here, can overcome these assumptions with increased robustness to skewness and outliers. The use of a false discovery rate provided a further method for conservative interpretation of these results. We combined *L. crispatus*-dominated CST-I and *L. iners*-dominated CST-III because in this study, succinate was the only metabolite with abundance significantly different between samples belonging to CST-I compared to CST-III. Combining the two CSTs allowed us to increase our statistical power by comparing CST-I/III to CST-IV, however combining the *L. crispatus* and *L. iners*-dominated communities may not be functionally optimal. There is a growing body of research focused on *L. iners*<sup>32,100–103</sup> that aims to evaluate its protective value as part of the vaginal microbiota. *L. iners* is commonly detected in healthy women, and interestingly in women with BV, as well as being among the first *Lactobacillus* spp. to recover after antibiotic treatment for BV<sup>104</sup>. Further, it is often considered to be a tipping point for some women at risk for BV<sup>104</sup>. Despite the small sample size, this study is unique in its use of metabolomics and microbial abundance data in combination with extensive participant behavioral data and its rigorous evaluation of smoking status by using quantification of carbon monoxide exhaled and cotinine in saliva. The depth of the analyses performed in this study, and the results obtained, may inform where to focus resources in a larger confirmatory study.

## Conclusions

It is well-documented that a vaginal microbiota dominated by *Lactobacillus* spp. is associated with reduced risk for urogenital infections, including sexually transmitted infections (STIs) and urinary tract (UTIs) infections<sup>105</sup>. In this study, we determined that overall smoking did not affect the vaginal metabolome after controlling for CST, but several key metabolites were elevated in smokers. Among women who were smokers and depauperate for *Lactobacillus* spp. (classified as CST-IV), we observed that they had significantly more perturbed metabolic profile than other CSTs when compared to their non-smoking counterparts. Biogenic amines were elevated in smokers and these metabolites have known roles in anaerobic bacterial proliferation, immune- and stress-resistance with a significant link to the development of BV, and possibly other reproductive tract infections. Women who smoke may have increased susceptibility to reproductive tract infections due to the observed increase in concentrations of BAs and the finding was even more pronounced among women who had low levels of *Lactobacillus* spp.

The metabolite profile of the vagina was strongly influenced by the resident microbiota as well as cigarette smoking in epidemiologic analyses that controlled for possible confounders. Detection of nicotine and its breakdown products in the vagina may serve as molecular biomarkers of smoking. Our results suggest that smoking is associated with several important metabolites present in the vagina that may have implications for women's health. This study serves as a pilot for the development of future studies of the mechanisms linking smoking to poor gynecologic and reproductive health outcomes.

## Methods

**Sample Collection.** Forty women self-collected mid-vaginal swabs during a single visit at the Center for Health Behavior Research at the University of Maryland School of Public Health (UMSPH). The study has previously been described<sup>30</sup>. In brief, smoking burden was determined by saliva cotinine testing, carbon monoxide exhalation levels and self-report of smoking habits on a set of comprehensive behavioral surveys. Participants were excluded if they had used an antibiotic or antimycotic in the 30 days prior. Four women were excluded from this analysis due to poor DNA quality affecting the normalization of the vaginal metabolome dataset (final sample size for analysis was 36 (17 smokers and 19 non-smokers) (Table 2). All participants provided written informed consent, and ethical approval was obtained from the Institutional Review Boards of the University of Maryland Baltimore (UMB), Montana State University and the UMSPH. All samples were collected and analysed in accordance with the relevant guidelines and regulations.

**Sample Preparation for Metabolomics.** Vaginal samples were eluted from swabs (Starplex rayon swab) in 200 µl phosphate buffered saline (PBS) and subjected to both gas chromatography mass spectrometry (GC-MS) and liquid chromatography mass spectrometry (LC-MS) with Orbitrap Elite accurate mass platforms (Thermo Scientific, Waltham, MA, USA). Sample processing and analysis was performed by Metabolon (Durham, NC, USA) using an automated MicroLab STAR<sup>®</sup> system (Hamilton Company, Reno, NV, USA). Recovery standards were added prior to the first step in the extraction process for quality control purposes. Sample preparation was conducted using a proprietary series of organic and aqueous extractions to remove the protein fraction while allowing maximum recovery of small molecules. The resulting extract was divided into two fractions: one for analysis by LC and one for analysis by GC. Samples were placed briefly on a TurboVap<sup>®</sup> (Zymark, Hopkinton, MA, USA) to remove the organic solvent. Each sample was then frozen and dried under vacuum. Samples were then prepared for the appropriate instrument, either LC-MS or GC-MS.

**Liquid Chromatography and Gas Chromatography Mass Spectrometry.** LC-MS measurements were conducted on a Waters ACQUITY ultra-performance liquid chromatograph (UPLC) and a ThermoFisher Scientific Orbitrap Elite high resolution/accurate mass spectrometer (Thermo Scientific, Waltham, MA, USA), which consisted of a heated electrospray ionization (HESI) source and orbitrap mass analyzer operated at a resolution 30,000 mass. The sample extract was dried then reconstituted in LC-compatible solvents, each of which contained eight or more injection standards at fixed concentrations to ensure injection and chromatographic consistency. One aliquot was analyzed using acidic positive ion optimized conditions and the other using basic negative ion optimized conditions in two independent injections using separate dedicated columns. Extracts

reconstituted in acidic conditions were gradient eluted using water and methanol containing 0.1% formic acid, while the basic extracts, which also used water/methanol, contained 6.5 mM ammonium bicarbonate. The MS analysis alternated between MS and data-dependent MS<sup>2</sup> scans using dynamic exclusion.

Samples for GC-MS analysis were re-dried under vacuum desiccation for a minimum of 24 hours prior to being derivatized under dried nitrogen using bistrimethyl-silyl-trifluoroacetamide (BSTFA). The GC column was 5% phenyl and the temperature ramp is from 40° to 300°C in a 16 min period. Samples were analyzed on a Thermo-Finnigan Trace DSQ fast-scanning single-quadrupole mass spectrometer using electron impact ionization. The instrument was tuned and calibrated for mass resolution and mass accuracy prior to use. Peaks in the GC-MS and LC-MS data were identified using Metabolon's proprietary peak integration software to resolve sample metabolite peaks over noise. Complete details of methods describing metabolomic profiling are described in Lawton *et al.* (2008)<sup>106</sup>.

**Compound Identification and Preparation.** Spectra corresponding to each metabolite were identified by comparison to library entries of purified metabolite standards and their distinction from more than 1,000 other commercially available purified standard compounds. The combination of chromatographic properties and mass spectra gave an indication of a match to the specific biogenic amine compound or an isobaric entity. Results were manually curated to ensure that data were accurate and to remove any system artifacts, miss assignments, and background noise.

**Taxonomic Assignment and Community State Type Profiling.** DNA extraction, PCR amplification and sequencing of 16S rRNA gene amplicons from the vaginal tract of study participants were conducted in a prior study<sup>30</sup>. Briefly, the V1-V3 region of the 16S rRNA gene was PCR amplified using the primers 27F-YM + 3<sup>107</sup> and 534R<sup>57</sup> and pyrosequenced using a Roche 454 FLX instrument. Species level assignments of *Lactobacillus* were performed using higher order Markov Chain models using the software speciateIT (speciateIT.sourceforge.net)<sup>33</sup>. For each sample, community state types (CSTs) were assigned to individual samples based on diversity and relative abundances of different phylotypes as defined in the work of Gajer and Brotman *et al.*<sup>108</sup>. The vaginal microbiota samples were categorized into CST-I (*L. crispatus*-dominated), CST-III (*L. iners*-dominated) and CST-IV (low-*Lactobacillus*/high strict and facultative anaerobes) (Table 2) Two other well-documented CSTs, CST-II (*L. gasseri*-dominated) and CST-V (*L. jensenii*-dominated) were not identified in this limited sample and are less commonly found even in larger surveys of women<sup>33</sup>.

**Data Analyses.** The metabolomic dataset was normalized to DNA concentration and missing values were imputed with the minimum detected value for that compound which essentially assigns values based on the sensitivity limit<sup>109</sup>. Data were then centered and scaled and the median was set equal to 1 prior to log transformation. Samples were visualized using a principal components analysis (PCA) (Figure S1). We performed quantile regression to estimate the median metabolite concentration differences 1) between smokers and non-smokers, 2) between CSTs and, 3) also with adjustment for confounding factors. P-values were obtained using the Wilcoxon rank-sum test. To correct for multiple tests, false discovery rate (FDR) and q-values were calculated for each compound where  $p = < 0.05$ . Where a p-value estimates the proportion of all tests which will result in false positives (i.e. 5% where  $p\text{-value} = 0.05$ ), a q-value estimates the proportion of significant tests that will result in false positives<sup>110</sup>. Initial investigations between *Lactobacillus*-dominated CST-I and CST-III yielded only one significant difference in the metabolite succinate ( $q\text{-value} = 0.009$ ) with adjustment for smoking status. This metabolite was not identified as differing significantly ( $q\text{-value} = 0.05$ ) in all further analyses, and therefore we grouped CST-I and CST-III for further binary analyses to increase power in comparisons to CST-IV. Fold change (FC) was calculated based on Guo *et al.*<sup>111</sup> where  $FC_i = x_i - y_i$  where  $i$  is metabolite of interest for the mean value of the control ( $x$ ) and treatment ( $y$ ).

We performed distance based linear modeling (DISTLM) by partitioning the distances of compounds using distance-based redundancy analysis (dbRDA)<sup>112</sup>, a form of multivariate multiple regression that can be performed directly on a distance or dissimilarity matrix. As the number of participants limits the number of variables that can be included in the model, we planned to include just two variables, in addition to our main factors of CST and smoking status. Initially we ran individual variables against the metabolite data and selected only those variables that were significant ( $p \leq 0.05$ ). These were age, race, marital status, education level, thong undergarment in the prior 24 hours, and vaginal douching in the two months prior. The dissimilarity matrix of vaginal metabolites was fitted with four of the variables listed in Table 2 and the final model was selected with the greatest adjusted R<sup>2</sup> value.

Boxplots and heatmaps were constructed using the ggplot2<sup>113</sup> and RColorBrewer (colorbrewer2.org) packages conducted in R version 3.1.2<sup>114</sup>. PCA, correlation coefficients, DISTLM and dbRDA plots were produced in Primer version 6<sup>115</sup>.

**Data availability.** The questionnaire and metabolome data are available at the National Center for Biotechnology Information (NCBI) Database of Genotypes and Phenotypes (dbGaP) under Accession number phs001386.v1.p1. Metagenomic sequence data were submitted to the public NCBI Sequence Read Archive (SRA) with the accession number PRJNA391039.

## References

1. National Center for Health Statistics. *Early Release of Selected Estimates Based on Data from the National Health Interview Survey, 2016* (2017).
2. Fact sheet: Current Cigarette Smoking Among Adults in the United States. *JAMA* **309** (2013).
3. Eschenbach, D. A. *Diagnosis of Bacterial Vaginosis (Nonspecific Vaginitis): Role of the Laborator*. **6**, 134–138 (1984).
4. Tripathi, R., Dimri, S., Bhalla, P. & Ramji, S. Bacterial vaginosis and pregnancy outcome. *International Journal of Gynecology & Obstetrics* **83**, 193–195 (2003).
5. van Oostrum, N., De Sutter, P., Meys, J. & Verstraelen, H. Risks associated with bacterial vaginosis in infertility patients: a systematic review and meta-analysis. *Human reproduction* **28**, 1809–1815 (2013).

6. Feldkamp, M. L. *et al.* Case-control study of self reported genitourinary infections and risk of gastroschisis: findings from the national birth defects prevention study, 1997–2003. *Bmj* **336**, 1420–1423 (2008).
7. Bagaitkar, J., Demuth, D. R. & Scott, D. A. *Tobacco use increases susceptibility to bacterial infection.* **4**, 12–10 (2008).
8. Larsson, P. G., Platz-Christensen, J. J. & Sundström, E. Is bacterial vaginosis a sexually transmitted disease? *International journal of STD & AIDS* **2**, 362–364 (1991).
9. Hellberg, D., Nilsson, S. & Mårdh, P. A. Bacterial vaginosis and smoking. *International journal of STD & AIDS* **11**, 603–606 (2000).
10. Hay, P. E. *et al.* A longitudinal study of bacterial vaginosis during pregnancy. *International Journal of Gynecology & Obstetrics* **3**, 320 (1995).
11. Nelson, D. B. *et al.* Vaginal symptoms and bacterial vaginosis (BV): how useful is self-report? Development of a screening tool for predicting BV status. *Epidemiol. Infect.* **135**, 1369–1375 (2007).
12. Miller, G. C., McDermott, R., McCulloch, B., Fairley, C. K. & Muller, R. Predictors of the prevalence of bacterial STI among young disadvantaged Indigenous people in north Queensland, Australia. *Sexually transmitted infections* **79**, 332–335 (2003).
13. Chernes, T. L., Hillier, S. L., Meyn, L. A., Busch, J. L. & Krohn, M. A. A delicate balance: risk factors for acquisition of bacterial vaginosis include sexual activity, absence of hydrogen peroxide-producing lactobacilli, black race, and positive herpes simplex virus type 2 serology. *Sex Transm Dis* **35**, 78–83 (2008).
14. Swartzendruber, A., Sales, J. M., Brown, J. L., DiClemente, R. J. & Rose, E. S. Correlates of Incident *Trichomonas vaginalis* Infections Among African American Female Adolescents. *Sexually transmitted diseases* **41**, 240–245 (2014).
15. Chernes, T. L., Meyn, L. A., Krohn, M. A. & Hillier, S. L. Risk factors for infection with herpes simplex virus type 2: Role of smoking, douching, uncircumcised males, and vaginal flora. *Sex Transm Dis* **30**, 405–410 (2003).
16. Wangnapi, R. A. *et al.* Prevalence and risk factors for Chlamydia trachomatis, Neisseria gonorrhoeae and *Trichomonas vaginalis* infection in pregnant women in Papua New Guinea. *Sexually transmitted infections* **91**, 194–+ (2015).
17. Marcone, V. *et al.* Epidemiology of Chlamydia trachomatis endocervical infection in a previously unscreened population in Rome, Italy, 2000 to 2009. **17**, 16–23 (2012).
18. Miller, C. S. & White, D. K. Human papillomavirus expression in oral mucosa, premalignant conditions, and squamous cell carcinoma: A retrospective review of the literature. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology* **82**, 57–68 (1996).
19. Gillison, M. L. *et al.* Distinct risk factor profiles for human papillomavirus type 16-positive and human papillomavirus type 16-negative head and neck cancers. **100**, 407–420 (2008).
20. Castellsagué, X. & Muñoz, N. Chapter 3: Cofactors in human papillomavirus carcinogenesis—role of parity, oral contraceptives, and tobacco smoking. *J. Natl. Cancer Inst. Monographs* 20–28 (2003).
21. Garrett, L. R., Perez-Reyes, N., Smith, P. P. & McDougall, J. K. Interaction of HPV-18 and nitrosomethylurea in the induction of squamous cell carcinoma. *Carcinogenesis* **14**, 329–332 (1993).
22. Nischan, P., Ebeling, K. & Schindler, C. Smoking and Invasive Cervical-Cancer Risk - Results From a Case-Control Study. **128**, 74–77 (1988).
23. Khan, A. M., Freeman-Wang, T., Pital, N. & Singer, A. Smoking and multicentric vulval intraepithelial neoplasia. *J Obstet Gynaecol* **29**, 123–125 (2009).
24. Arnon, Y., Shoenfeld, Y. & Amital, H. Effects of tobacco smoke on immunity, inflammation and autoimmunity. *J. Autoimmun.* **34**, J258–J265 (2010).
25. Simons, A. M., Phillips, D. H. & Coleman, D. V. Damage to DNA in Cervical Epithelium Related to Smoking. *Tobacco*. **49**, 186 (1994).
26. Bornstein, J., Rahat, M. A. & Abramovici, H. Etiology of Cervical Cancer: Current Concepts. **50**, 146 (1995).
27. McCann, M. F. *et al.* Nicotine and cotinine in the cervical mucus of smokers, passive smokers, and nonsmokers. **1**, 125–129 (1992).
28. Winkelstein, W. Smoking and Cervical-Cancer - Current Status - a Review. **131**, 945–957 (1990).
29. Biedermann, L. *et al.* Smoking cessation induces profound changes in the composition of the intestinal microbiota in humans. **8**, e59260 (2013).
30. Brotman, R. M. *et al.* Association between cigarette smoking and the vaginal microbiota: a pilot study. *BMC Infect. Dis.* **14**, 471 (2014).
31. O'Hanlon, D. E., Moench, T. R. & Cone, R. A. Vaginal pH and Microbicidal Lactic Acid When Lactobacilli Dominate the Microbiota. *PLoS ONE* **8**, ARTN e80074 (2013).
32. Macklaim, J. M., Gloor, G. B., Anukam, K. C., Cribby, S. & Reid, G. At the crossroads of vaginal health and disease, the genome sequence of *Lactobacillus iners* AB-1. *Proc. Natl. Acad. Sci. USA* **108**, 4688–4695 (2011).
33. Ravel, J. *et al.* Vaginal microbiome of reproductive-age women. *Proc. Natl. Acad. Sci. USA* **108**(Suppl 1), 4680–4687 (2011).
34. Mehta, S. D. *et al.* The Vaginal Microbiota over an 8- to 10-Year Period in a Cohort of HIV-Infected and HIV-Uninfected Women. *PLoS ONE* **10**, e0116894 (2015).
35. Albert, A. Y. K. *et al.* A Study of the Vaginal Microbiome in Healthy Canadian Women Utilizing cpn 60-Based Molecular Profiling Reveals Distinct *Gardnerella* Subgroup Community State Types. **10**, e0135620 (2015).
36. Srinivasan, S. *et al.* More Than Meets the Eye: Associations of Vaginal Bacteria with Gram Stain Morphotypes Using Molecular Phylogenetic Analysis. **8** (2013).
37. Chernes, T. L., Meyn, L. A., Krohn, M. A., Lurie, J. G. & Hillier, S. L. Association between acquisition of herpes simplex virus type 2 in women and bacterial vaginosis. *Clin Infect Dis* **37**, 319–325 (2003).
38. Martin, H. L. *et al.* Vaginal lactobacilli, microbial flora, and risk of human immunodeficiency virus type 1 and sexually transmitted disease acquisition. *J Infect Dis* **180**, 1863–1868 (1999).
39. 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).
40. King, C. C. *et al.* Bacterial vaginosis and the natural history of human papillomavirus. *Infectious Diseases in Obstetrics and Gynecology* **2011**, 319460–8 (2011).
41. Gallo, M. F. *et al.* Bacterial Vaginosis, Gonorrhoea, and Chlamydial Infection Among Women Attending a Sexually Transmitted Disease Clinic: A Longitudinal Analysis of Possible Causal Links. *Annals of epidemiology* **22**, 213–220 (2012).
42. Balkus, J. E. *et al.* Bacterial vaginosis and the risk of *Trichomonas vaginalis* acquisition among HIV-1-negative women. *Sexually transmitted diseases* **41**, 123–128 (2014).
43. Ghartey, J. P. *et al.* *Lactobacillus crispatus* Dominant Vaginal Microbiome Is Associated with Inhibitory Activity of Female Genital Tract Secretions against *Escherichia coli*. **9** (2014).
44. Klupczynska, A., Dereziński, P. & Kokot, Z. J. Metabolomics in Medical Sciences - Trends, Challenges and Perspectives. *Acta Polonae Pharmaceutica* **72**, 629–641 (2015).
45. Yeoman, C. J. *et al.* A multi-omic systems-based approach reveals metabolic markers of bacterial vaginosis and insight into the disease. **8**, e56111 (2013).
46. Srinivasan, S. *et al.* Bacterial communities in women with bacterial vaginosis: high resolution phylogenetic analyses reveal relationships of microbiota to clinical criteria. *PLoS ONE* **7**, e37818 (2012).
47. McMillan, A. *et al.* A multi-platform metabolomics approach identifies highly specific biomarkers of bacterial diversity in the vagina of pregnant and non-pregnant women. *Scientific Reports* **5**, 14174 (2015).

48. Vitali, B. *et al.* Vaginal microbiome and metabolome highlight specific signatures of bacterial vaginosis. *Eur J Clin Microbiol* 1–10 <https://doi.org/10.1007/s10096-015-2490-y> (2015).
49. Aroutcheva, A. A., Simoes, J. A. & Faro, S. Antimicrobial protein produced by vaginal *Lactobacillus acidophilus* that inhibits *Gardnerella vaginalis*. *Infectious Diseases in Obstetrics and Gynecology* **9**, 33–39 (2001).
50. Turovskiy, Y. *et al.* Susceptibility of *Gardnerella vaginalis* biofilms to natural antimicrobials subtilisin, epsilon-poly-L-lysine, and lauramide arginine ethyl ester. *Infectious Diseases in Obstetrics and Gynecology* **2012**, 284762 (2012).
51. Braga, P. C., Dal Sasso, M., Culici, M. & Spallino, A. Inhibitory activity of thymol on native and mature *Gardnerella vaginalis* biofilms: *in vitro* study. *Arzneimittel-Forschung* **60**, 675–681 (2010).
52. Kanjee, U. & Houry, W. A. Mechanisms of acid resistance in *Escherichia coli*. **67**, 65–81 (2013).
53. Nasrallah, G. K. *et al.* *Legionella pneumophila* requires polyamines for optimal intracellular growth. **193**, 4346–4360 (2011).
54. Jelsbak, L., Thomsen, L. E., Wallrodt, I., Jensen, P. R. & Olsen, J. E. Polyamines are required for virulence in *Salmonella enterica* serovar Typhimurium. *PLoS ONE* **7**, e36149 (2012).
55. Goytia, M. & Shafer, W. M. Polyamines can increase resistance of *Neisseria gonorrhoeae* to mediators of the innate human host defense. *Infection and immunity* **78**, 3187–3195 (2010).
56. Nelson, T. M. *et al.* Vaginal biogenic amines: biomarkers of bacterial vaginosis or precursors to vaginal dysbiosis? **6**, 253 (2015).
57. Ravel, J. *et al.* Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. *Microbiome* **1**, 1 (2013).
58. Hukkanen, J., Jacob, P. & Benowitz, N. L. Metabolism and disposition kinetics of nicotine. *Pharmacol. Rev.* **57**, 79–115 (2005).
59. Baheiraei, A., Banihosseini, S. Z., Heshmat, R., Mota, A. & Mohsenifar, A. Association of self-reported passive smoking in pregnant women with cotinine level of maternal urine and umbilical cord blood at delivery. *Paediatr Perinat Epidemiol* **26**, 70–76 (2012).
60. Dempsey, D., Jacob, P. & Benowitz, N. L. Accelerated metabolism of nicotine and cotinine in pregnant smokers. *The Journal of pharmacology and experimental therapeutics* **301**, 594–598 (2002).
61. Benowitz, N. L. Cotinine as a biomarker of environmental tobacco smoke exposure. *Epidemiol Rev* **18**, 188–204 (1996).
62. Poppe, W. A., Peeters, R., Daenens, P., Ide, P. S. & Vanassche, F. A. Tobacco Smoking and the Uterine Cervix - Cotinine in. *Blood, Urine and Cervical Fluid*. **39**, 110–114 (1995).
63. Schiffman, M. H. *et al.* Biochemical epidemiology of cervical neoplasia: measuring cigarette smoke constituents in the cervix. *Cancer Res* **47**, 3886–3888 (1987).
64. Mietz, J. L. & Karmas, E. Polyamine and histamine content of rockfish, salmon, lobster, and shrimp as an indicator of decomposition. *Journal of the Association of Official Analytical Chemists (USA)* (1978).
65. Hussain, A. *et al.* High-affinity olfactory receptor for the death-associated odor cadaverine. *Proc. Natl. Acad. Sci. USA* **110**, 19579–19584 (2013).
66. Chen, K. C. *et al.* Amine content of vaginal fluid from untreated and treated patients with nonspecific vaginitis. *The Journal of clinical investigation* **63**, 828–835 (1979).
67. Sobel, J. D. *et al.* Diagnosing vaginal infections through measurement of biogenic amines by ion mobility spectrometry. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **163**, 81–84 (2012).
68. Benowitz, N. L. Clinical Pharmacology of Nicotine: Implications for Understanding, Preventing, and Treating Tobacco. *Addiction*. **83**, 531–541 (2008).
69. Strom, A. R. *et al.* Trimethylamine oxide: a terminal electron acceptor in anaerobic respiration of bacteria. *Journal of general microbiology* **112**, 315–320 (1979).
70. O'Hanlon, D. E., Moench, T. R. & Cone, R. A. In vaginal fluid, bacteria associated with bacterial vaginosis can be suppressed with lactic acid but not hydrogen peroxide. *BMC infectious diseases* **11**, 200 (2011).
71. Koumans, E. H. *et al.* The prevalence of bacterial vaginosis in the United States, 2001–2004; associations with symptoms, sexual behaviors, and reproductive health. *Sexually transmitted diseases* **34**, 864–869 (2007).
72. Li, P., Beck, W. D., Callahan, P. M., Terry, A. V. J. & Bartlett, M. G. Quantitation of cotinine and its metabolites in rat plasma and brain tissue by hydrophilic interaction chromatography tandem mass spectrometry (HILIC-MS/MS). *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **907**, 117–125 (2012).
73. Brunnemann, K. D., Kagan, M. R., Cox, J. E. & Hoffmann, D. Chemical Studies on Tobacco-Smoke.87. Determination of Benzene, Toluene and 1,3-Butadiene in Cigarette-Smoke by Gc-Msd. *Exp Pathol* **37**, 108–113 (1989).
74. Piot, P., Van Dyck, E., Totten, P. A. & Holmes, K. K. Identification of *Gardnerella (Haemophilus) vaginalis*. **15**, 19–24 (1982).
75. Aroutcheva, A. A., Simoes, J. A., Behbakht, K. & Faro, S. *Gardnerella vaginalis* isolated from patients with bacterial vaginosis and from patients with healthy vaginal ecosystems. **33**, 1022–1027 (2001).
76. Spiegel, C. A. *et al.* *Gardnerella-Vaginalis* and Anaerobic-Bacteria in the Etiology of Bacterial (Nonspecific) Vaginosis. 41–46 (1983).
77. Swidsinski, A. *et al.* An adherent *Gardnerella vaginalis* biofilm persists on the vaginal epithelium after standard therapy with oral metronidazole. *Am J Obstet Gynecol* **198**, 97 e1–6 (2008).
78. Leitich, H. *et al.* Bacterial vaginosis as a risk factor for preterm delivery: A meta-analysis. *Am J Obstet Gynecol* **189**, 139–147 (2003).
79. Delbridge, L. M. & O'Riordan, M. X. Innate recognition of intracellular bacteria. *Current opinion in immunology* **19**, 10–16 (2007).
80. Strom, K., Sjogren, J., Broberg, A. & Schnurer, J. *Lactobacillus plantarum* MiLAB 393 Produces the Antifungal Cyclic Dipeptides *Cyclo(L-Phe-L-Pro)* and *Cyclo(L-Phe-trans-4-OH-L-Pro)* and 3-Phenyllactic Acid. **68**, 4322–4327 (2002).
81. Piuri, M., Sanchez-Rivas, C. & Ruzal, S. M. Adaptation to high salt in *Lactobacillus*: role of peptides and proteolytic enzymes. *J. Appl. Microbiol.* **95**, 372–379 (2003).
82. Magnusson, J., Ström, K., Roos, S., Sjögren, J. & Schnürer, J. Broad and complex antifungal activity among environmental isolates of lactic acid bacteria. *FEMS Microbiol. Lett.* **219**, 129–135 (2003).
83. Houston, D. R. *et al.* Structure-based exploration of cyclic dipeptide chitinase inhibitors. *J Med Chem* **47**, 5713–5720 (2004).
84. Yang, E. J. & Chang, H. C. Purification of a new antifungal compound produced by *Lactobacillus plantarum* AF1 isolated from kimchi. *Int J Food Microbiol* **139**, 56–63 (2010).
85. Sinha, S., Srivastava, R., De Clercq, E. & Singh, R. K. Synthesis and Antiviral Properties of Arabino and Ribonucleosides of 1,3-Dideazaadenine, 4-Nitro-1, 3-dideazaadenine and Diketopiperazine. *Nucleosides, Nucleotides & Nucleic Acids* **23**, 1815–1824 (2004).
86. Kwon, O. S., Park, S. H., Yun, B. S., Pyun, Y. R. & Kim, C. J. *Cyclo(Dehydroala-L-Leu)*, an alpha-glucosidase inhibitor from *Penicillium* sp F70614. **53**, 954–958 (2000).
87. Olmsted, S. S., Meyn, L. A., Rohan, L. C. & Hillier, S. L. Glycosidase and proteinase activity of anaerobic gram-negative bacteria isolated from women with bacterial vaginosis. *Sex Transm Dis* **30**, 257–261 (2003).
88. Wiggins, R., Hicks, S. J., Soothill, P. W., Millar, M. R. & Corfield, A. P. Mucinases and sialidases: their role in the pathogenesis of sexually transmitted infections in the female genital tract. *Sexually transmitted infections* **77**, 402–408 (2001).
89. Briselden, A. M., Moncla, B. J., Stevens, C. E. & Hillier, S. L. Sialidases (neuraminidases) in bacterial vaginosis and bacterial vaginosis-associated microflora. *J. Clin. Microbiol.* **30**, 663–666 (1992).
90. McGregor, J. A. *et al.* Bacterial vaginosis is associated with prematurity and vaginal fluid mucinase and sialidase: Results of a controlled trial of topical clindamycin cream. *Am J Obstet Gynecol* **170**, 1048–1060 (1994).
91. Straus, D. C., Mattingly, S. J., Milligan, T. W., Doran, T. I. & Nealon, T. J. Protease production by clinical isolates of type III group B streptococci. *J. Clin. Microbiol.* **12**, 421–423 (1980).



92. Volpe, G. E. *et al.* Associations of cocaine use and HIV infection with the intestinal microbiota, microbial translocation, and inflammation. *J Stud Alcohol Drugs* **75**, 347–357 (2014).
93. Khan, M. R. *et al.* Non-injection and injection drug use and STI/HIV risk in the United States: the degree to which sexual risk behaviors versus sex with an STI-infected partner account for infection transmission among drug users. *AIDS Behav* **17**, 1185–1194 (2013).
94. Cavazos-Rehg, P. A. *et al.* Risky sexual behaviors and sexually transmitted diseases: a comparison study of cocaine-dependent individuals in treatment versus a community-matched sample. *Aids Patient Care St* **23**, 727–734 (2009).
95. Uguz, F. *et al.* Antidepressants and menstruation disorders in women: a cross-sectional study in three centers. *Gen Hosp Psychiatry* **34**, 529–533 (2012).
96. Urban, R. J. & Veldhuis, J. D. A selective serotonin reuptake inhibitor, fluoxetine hydrochloride, modulates the pulsatile release of prolactin in postmenopausal women. **164**, 147–152 (1991).
97. Srinivasan, S. *et al.* Metabolic signatures of bacterial vaginosis. **6** (2015).
98. Gika, H. G., Theodoridis, G. A., Plumb, R. S. & Wilson, I. D. Current practice of liquid chromatography-mass spectrometry in metabolomics and metabonomics. *Journal of Pharmaceutical and Biomedical Analysis* **87**, 12–25 (2014).
99. Courant, F., Antignac, J.-P., Dervilly-Pinel, G. & Le Bizet, B. Basics of mass spectrometry based metabolomics. **14**, 2369–2388 (2014).
100. Fredricks, D. N. Molecular methods to describe the spectrum and dynamics of the vaginal microbiota. *Anaerobe* **17**, 191–195 (2011).
101. Verstraelen, H. *et al.* Longitudinal analysis of the vaginal microflora in pregnancy suggests that *L. crispatus* promotes the stability of the normal vaginal microflora and that *L. gasseri* and/or *L. iners* are more conducive to the occurrence of abnormal vaginal microflora. *BMC Microbiol.* **9**, 116 (2009).
102. Witkin, S. S. *et al.* Influence of vaginal bacteria and D- and L-lactic acid isomers on vaginal extracellular matrix metalloproteinase inducer: implications for protection against upper genital tract infections. **4** (2013).
103. McMillan, A., Macklaim, J. M., Burton, J. P. & Reid, G. Adhesion of *Lactobacillus iners* AB-1 to human fibronectin: a key mediator for persistence in the vagina? *Reprod Sci* **20**, 791–796 (2013).
104. Mayer, B. T. *et al.* Rapid and profound shifts in the vaginal microbiota following antibiotic treatment for bacterial vaginosis. *J Infect Dis* **212**, 793–802 (2015).
105. Martin, D. H. The microbiota of the vagina and its influence on women's health and disease. **343**, 2–9 (2012).
106. Lawton, K. A. *et al.* Analysis of the adult human plasma metabolome. <https://doi.org/10.2217/14622416.9.4.383> **9**, 383–397 (2008).
107. Frank, J. A. *et al.* Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. *Applied and Environmental Microbiology* **74**, 2461–2470 (2008).
108. Gajer, P. *et al.* Temporal Dynamics of the Human Vaginal Microbiota. *Sci Transl Med* **4**, 132ra52–132ra52 (2012).
109. Xia, J. & Wishart, D. S. Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nature protocols* **6**, 743–760 (2011).
110. Vinaixa, M. *et al.* A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. *Metabolites* **2**, 775–795 (2012).
111. Guo, L. *et al.* Rat toxicogenomic study reveals analytical consistency across microarray platforms. *Nature biotechnology* **24**, 1162–1169 (2006).
112. Legendre, P. & Anderson, M. J. Distance-Based Redundancy Analysis: Testing Multispecies Responses in Multifactorial Ecological Experiments. **69**, 1 (1999).
113. Warnes, G. R. *et al.* gplots: Various R programming tools for plotting data. *R package version 2* (2009).
114. Ihaka, R. & Gentleman, R. R. a language for data analysis and graphics. **5**, 299–314 (1996).
115. Clarke, K. & Gorley, R. *PRIMER v6: User manual/tutorial: PRIMER E.* (Plymouth) 2006.

## Acknowledgements

This project was funded by the National Institute of Allergy and Infectious Diseases (NIAID) grants K01-AI080974 (Brotman), R21-AI111145 (Yeoman), NR015495 (Ravel) the University of Maryland Cancer Epidemiology Alliance Joint Research Pilot Grant sponsored by the University of Maryland Greenebaum Cancer Center (Brotman), and the University of Maryland College Park / University of Maryland Baltimore Seed Grant (Brotman and Glover).

## Author Contributions

R.M.B., J.R., J.M.R. and E.G. designed and implemented the study. T.N., J.B., M.S., R.M.B., C.Y., R.M. and D.R. performed analyses of samples and/or data. T.N., R.M.B. and C.Y. wrote the main manuscript text. All authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-017-14943-3>.

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

**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) 2018