Inventory of Supplemental Information

1. Supplemental Figures

2. Supplemental Tables

3. Supplemental Experimental Procedures

4. Supplemental References

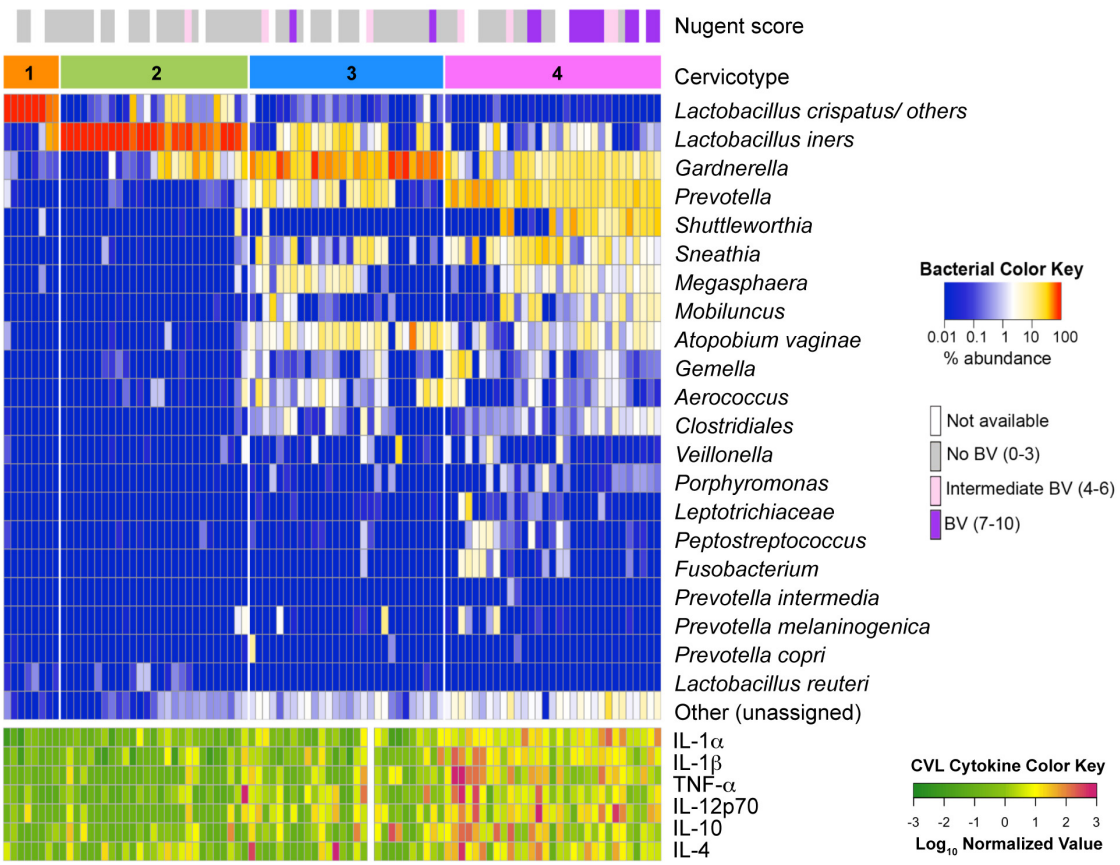


Figure S1, Related to Figure 1. Heatmap of bacterial taxa identified by 16S V4 sequencing (same as **Figure 1A**) with a second heatmap representing matched CVL cytokines: IL-1a, IL-1b, TNF-a, IL-12p70, IL-10, and IL-4. Cytokine concentrations were normalized as in **Figure 3A**.

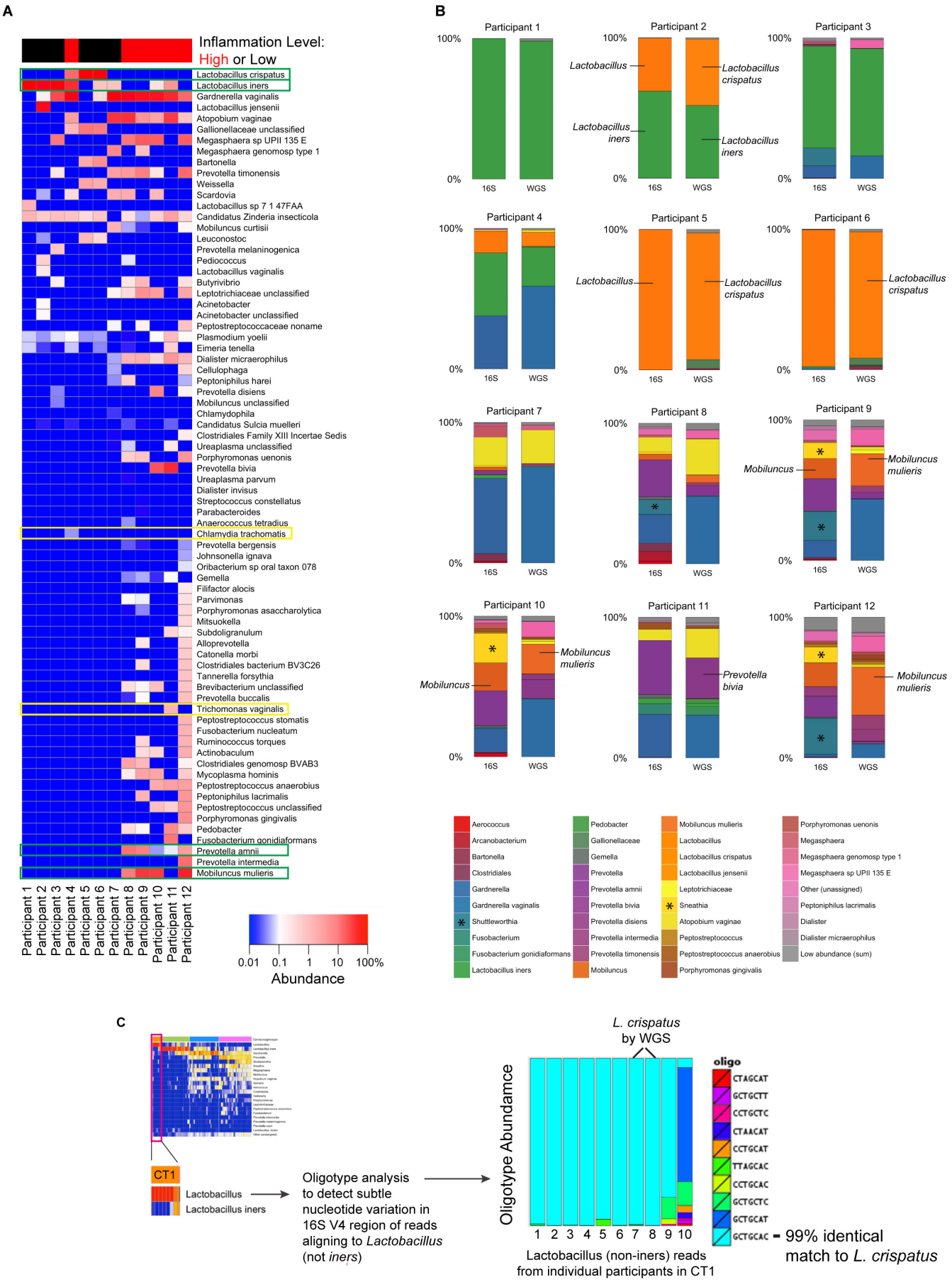

A
Inflammation Level:
High or Low
Lactobacillus crispatus
Lactobacillus iners
Gardnerella vaginalis
Lactobacillus jensenii
Atopobium vaginae
Gallionellaceae unclassified
Megasphaera sp UPII 135 E
Megasphaera genomosp type 1
Bartonella
Prevotella timonensis
Weissella
Scardovia
Lactobacillus sp 7 1 47FAA
Candidatus Zinderia insecticola
Mobiluncus curtisii
Leuconostoc
Prevotella melaninogenica
Pediococcus
Lactobacillus vaginalis
Butyrivibrio
Leptotrichiaceae unclassified
Acinetobacter
Acinetobacter unclassified
Peptostreptococcaceae noname
Plasmodium yoelii
Eimeria tenella
Dialister micraerophilus
Cellulophaga
Peptoniphilus harei
Prevotella disiens
Mobiluncus unclassified
Chlamydophila
Candidatus Sulcia muelleri
Clostridiales Family XIII Incertae Sedis
Ureaplasma unclassified
Porphyromonas uenonis
Prevotella bivia
Ureaplasma parvum
Dialister invisus
Streptococcus constellatus
Parabacteroides
Anaerococcus tetradius
Chlamydia trachomatis
Prevotella bergensis
Johnsonella ignava
Oribacterium sp oral taxon 078
Gemella
Filifactor alocis
Parvimonas
Porphyromonas asaccharolytica
Mitsuokella
Subdoligranulum
Alloprevotella
Catonella morbi
Clostridiales bacterium BV3C26
Tannerella forsythia
Brevibacterium unclassified
Prevotella buccalis
Trichomonas vaginalis
Peptostreptococcus stomatis
Fusobacterium nucleatum
Ruminococcus torques
Actinobaculum
Clostridiales genomosp BVAB3
Mycoplasma hominis
Peptostreptococcus anaerobius
Peptoniphilus lacrimalis
Peptostreptococcus unclassified
Porphyromonas gingivalis
Pedobacter
Fusobacterium gonidiaformans
Prevotella amnii
Prevotella intermedia
Mobiluncus mulieris
Participant 1
Participant 2
Participant 3
Participant 4
Participant 5
Participant 6
Participant 7
Participant 8
Participant 9
Participant 10
Participant 11
Participant 12
0.01 0.1 1 10 100%
Abundance
B
Participant 1
Participant 2
Participant 3
Participant 4
Participant 5
Participant 6
Participant 7
Participant 8
Participant 9
Participant 10
Participant 11
Participant 12
100%
0%
16S
WGS
Lactobacillus
Lactobacillus crispatus
Lactobacillus iners
Mobiluncus
Mobiluncus mulieris
Prevotella bivia
Aerococcus
Arcanobacterium
Bartonella
Clostridiales
Gardnerella
Gardnerella vaginalis
Shuttleworthia
Fusobacterium
Fusobacterium gonidiaformans
Lactobacillus iners
Pedobacter
Gallionellaceae
Gemella
Prevotella
Prevotella amnii
Prevotella bivia
Prevotella disiens
Prevotella intermedia
Prevotella timonensis
Mobiluncus
Mobiluncus mulieris
Lactobacillus
Lactobacillus crispatus
Lactobacillus jensenii
Leptotrichiaceae
Sneathia
Atopobium vaginae
Peptostreptococcus
Peptostreptococcus anaerobius
Porphyromonas gingivalis
Porphyromonas uenonis
Megasphaera
Megasphaera genomosp type 1
Megasphaera sp UPII 135 E
Other (unassigned)
Peptoniphilus lacrimalis
Dialister
Dialister micraerophilus
Low abundance (sum)
C
CT1
Lactobacillus
Lactobacillus iners
Oligotype analysis to detect subtle nucleotide variation in 16S V4 region of reads aligning to Lactobacillus (not iners)
L. crispatus by WGS
Oligotype Abundamce
1 2 3 4 5 6 7 8 9 10
Lactobacillus (non-iners) reads from individual participants in CT1
oligo
CTAGCAT
GCTGCTT
CCTGCTC
CTAACAT
CCTGCAT
TTAGCAC
CCTGCAC
GCTGCTC
GCTGCAT
GCTGCAC
= 99% identical match to L. crispatus

Figure S2, Related to Figure 2. Summary of metagenomic taxonomic identification using whole-genome shotgun sequencing and oligotyping.

(**A)** Heatmap of metagenomic taxonomic identification using the MetaPhlAn2 tool, with median clustering (dendrogram not shown). Six women with the highest pro-inflammatory cytokine levels (red) and six with the lowest levels (black) were chosen for metagenomic profiling. Taxa with abundances greater than 0.01% in any sample were shown. *Chlamydia trachomatis* and *Trichomonas vaginalis* were only detected in women who were PCR positive (yellow box). Species tested *in vitro* are highlighted in green.

(**B**) Comparison of 16S (left) and whole-genome shotgun sequencing (WGS; right) taxonomic identification for all twelve participants for whom both methods were used. Note that MetaPhlAn2, the program that we used for bacterial classification of the WGS reads, does not include *Sneathia* or *Shuttleworthia* in its reference database (due to the absence of full genome sequences for these bacteria), and thus cannot assign those genera. *Sneathia* or *Shuttleworthia* reads from 16S sequencing are marked with an asterisk.

(**C**) Oligotyping was performed on the reads that QIIME assigned to *Lactobacillus* (not *iners*) from women in CT1.

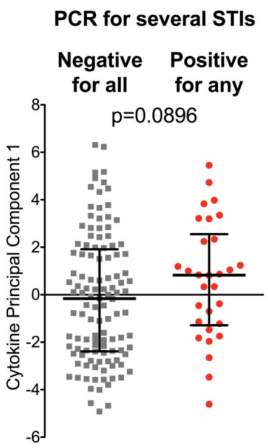
A

	IL-1RA	IP-10	RANTES	MIP-1α	IL-12p70	IL-10	TNF-α	IL-1β	IFN-γ	IL-2	IL-4	FLT-3L	IL-1α	MIP-1β	IFNα2	sCD40L
IL-1RA	1	0.37										0.32	0.31			
IP-10	0.37	1	0.53									0.34		0.37		
RANTES		0.53	1	0.42		0.4	0.32	0.31		0.38	0.31	0.34		0.58	0.38	0.48
MIP-1α			0.42	1	0.33	0.54	0.46	0.47	0.37	0.33	0.34			0.53		0.46
IL-12p70				0.33	1	0.61	0.72	0.5	0.53	0.6	0.42	0.32	0.4	0.38	0.31	0.47
IL-10			0.4	0.54	0.61	1	0.78	0.56	0.48	0.54	0.61			0.44		0.44
TNF-α			0.32	0.46	0.72	0.78	1	0.69	0.55	0.52	0.49		0.46	0.41	0.31	0.36
IL-1β			0.31	0.47	0.5	0.56	0.69	1	0.66	0.49	0.59	0.42	0.61	0.39		0.33
IFN-γ				0.37	0.53	0.48	0.55	0.66	1	0.54	0.63	0.46	0.58	0.39	0.31	0.36
IL-2			0.38	0.33	0.6	0.54	0.52	0.49	0.54	1	0.65	0.45	0.35	0.44	0.45	0.57
IL-4			0.31	0.34	0.42	0.61	0.49	0.59	0.63	0.65	1	0.44	0.43	0.34	0.28	0.47
FLT-3L	0.32	0.34	0.34		0.32			0.42	0.46	0.45	0.44	1	0.6	0.43	0.44	0.52
IL-1α	0.31				0.4		0.46	0.61	0.58	0.35	0.43	0.6	1	0.31	0.31	
MIP-1β		0.37	0.58	0.53	0.38	0.44	0.41	0.39	0.39	0.44	0.34	0.43	0.31	1	0.61	0.6
IFNα2			0.38		0.31		0.31		0.31	0.45	0.28	0.44	0.31	0.61	1	0.67
sCD40L			0.48	0.46	0.47	0.44	0.36	0.33	0.36	0.57	0.47	0.52		0.6	0.67	1

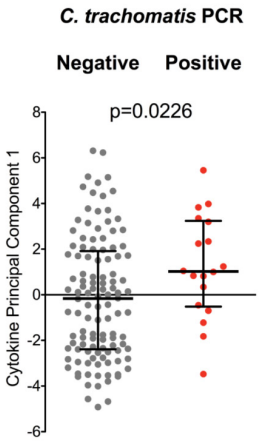
B

Loading of each cytokine	Principal Component	
	PC1 (41% of variation explained)	**PC2** (13% of variation explained)
TNF-α	0.303	-0.303
IL-1β	0.288	-0.213
IL-4	0.287	-0.114
IFN-γ	0.286	-0.088
IL-12p70	0.284	-0.202
IL-10	0.282	-0.255
IL-2	0.281	-0.061
sCD40L	0.269	0.176
MIP-1β	0.267	0.243
MIP-1α	0.244	-0.005
IL-1α	0.234	0.053
FLT-3L	0.232	0.312
IFN-α2	0.225	0.223
RANTES	0.216	0.199
IP-10	0.106	0.457
IL-1RA	0.040	0.494

C



D



E

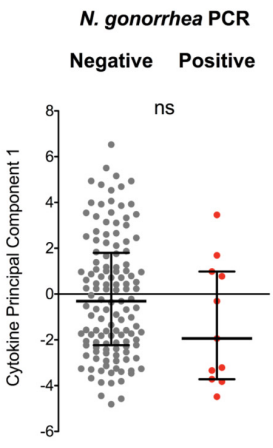


Figure S3, Related to Figure 3. Genital cytokine correlations.

(A) Spearman's rho correlation coefficients between CVL cytokines (using log10 normalized concentrations). Only correlations with $p<0.001$ are displayed.

(**B**) Loading of each cytokine on the first principal component (explaining 41% of variation) and the second principal component (explaining 13% of variation.)

(C-E) Genital cytokine levels in women with or without at least one positive STI PCR result (*C. trachomatis, N. gonorrhoea, T. vaginalis,* HSV-1, or HSV-2) (**C**), or positive only for *C. trachomatis* versus negative for all STIs (**D**), or positive only for *N. gonorrhoea* versus negative for all STIs (**E**). Data shown as median with interquartile range. P values determined by a two-tailed Mann-Whitney test, n=146.

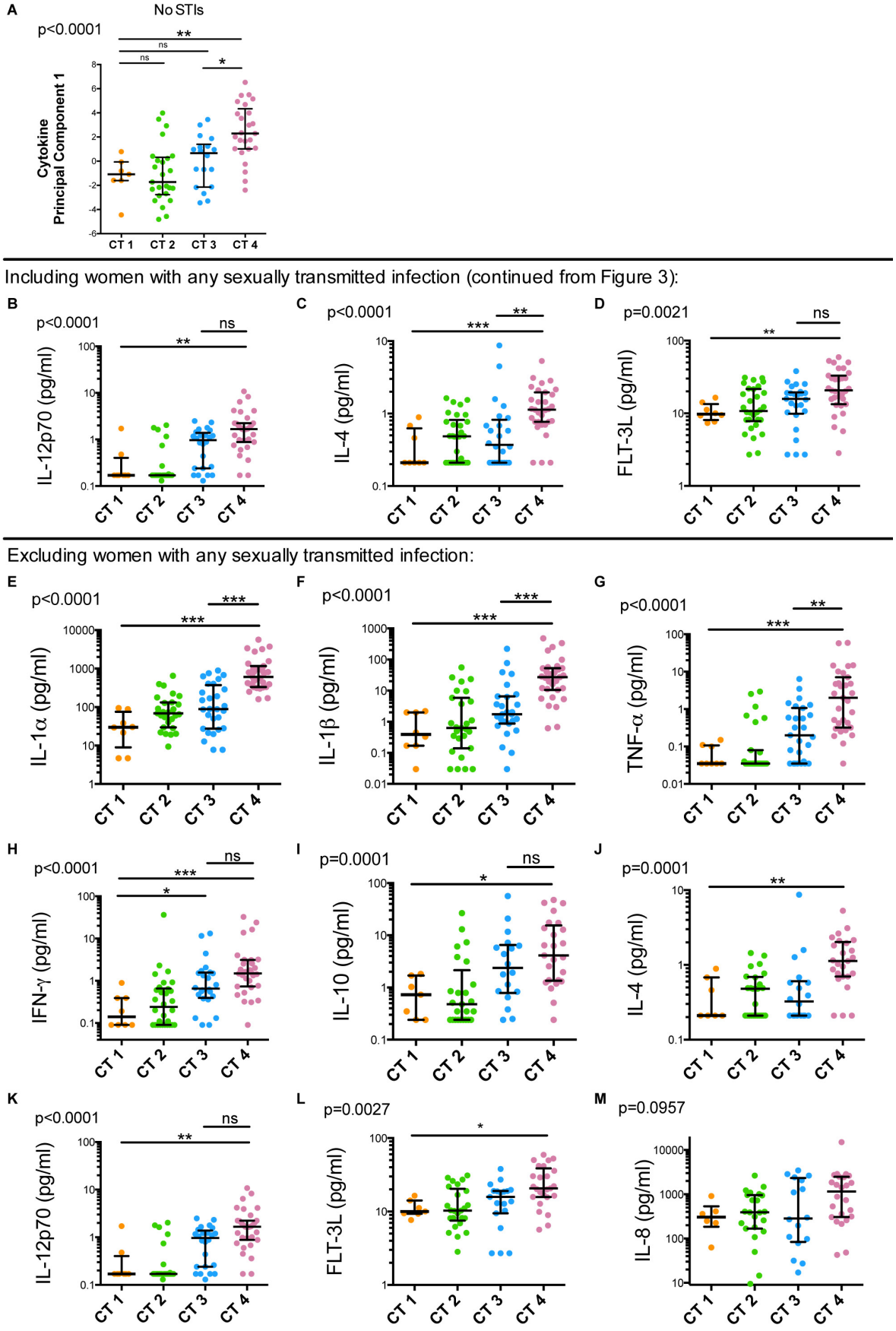

A
No STIs
p<0.0001
**
ns
ns
*
Cytokine Principal Component 1
CT 1
CT 2
CT 3
CT 4
Including women with any sexually transmitted infection (continued from Figure 3):
B
p<0.0001
ns
**
IL-12p70 (pg/ml)
C
p<0.0001
**
***
IL-4 (pg/ml)
D
p=0.0021
ns
**
FLT-3L (pg/ml)
Excluding women with any sexually transmitted infection:
E
p<0.0001
***
***
IL-1α (pg/ml)
F
p<0.0001
***
***
IL-1β (pg/ml)
G
p<0.0001
**
***
TNF-α (pg/ml)
H
p<0.0001
ns
***
*
IFN-γ (pg/ml)
I
p=0.0001
ns
*
IL-10 (pg/ml)
J
p=0.0001
**
IL-4 (pg/ml)
K
p<0.0001
ns
**
IL-12p70 (pg/ml)
L
p=0.0027
*
FLT-3L (pg/ml)
M
p=0.0957
IL-8 (pg/ml)

Figure S4, Related to Figure 4.

(**A**) Cytokine PC1 values from women with bacterial communities CT1-4, excluding women with any sexually transmitted infection (STI).

(**B-M**) Additional genital cytokines that are higher in CT4 than CT1, including women with STIs (**B-D**) or excluding women with any STIs (**E-M**). P values in upper left corners were determined by a Kruskal-Wallis test, and asterisks denote significance level after post-test. Data shown as median with interquartile range. * denotes $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

A

From \ To	CT1	CT2	CT3	CT4
CT1	1	4		
CT2	2	7	1	2
CT3		1	1	3
CT4		1	3	15

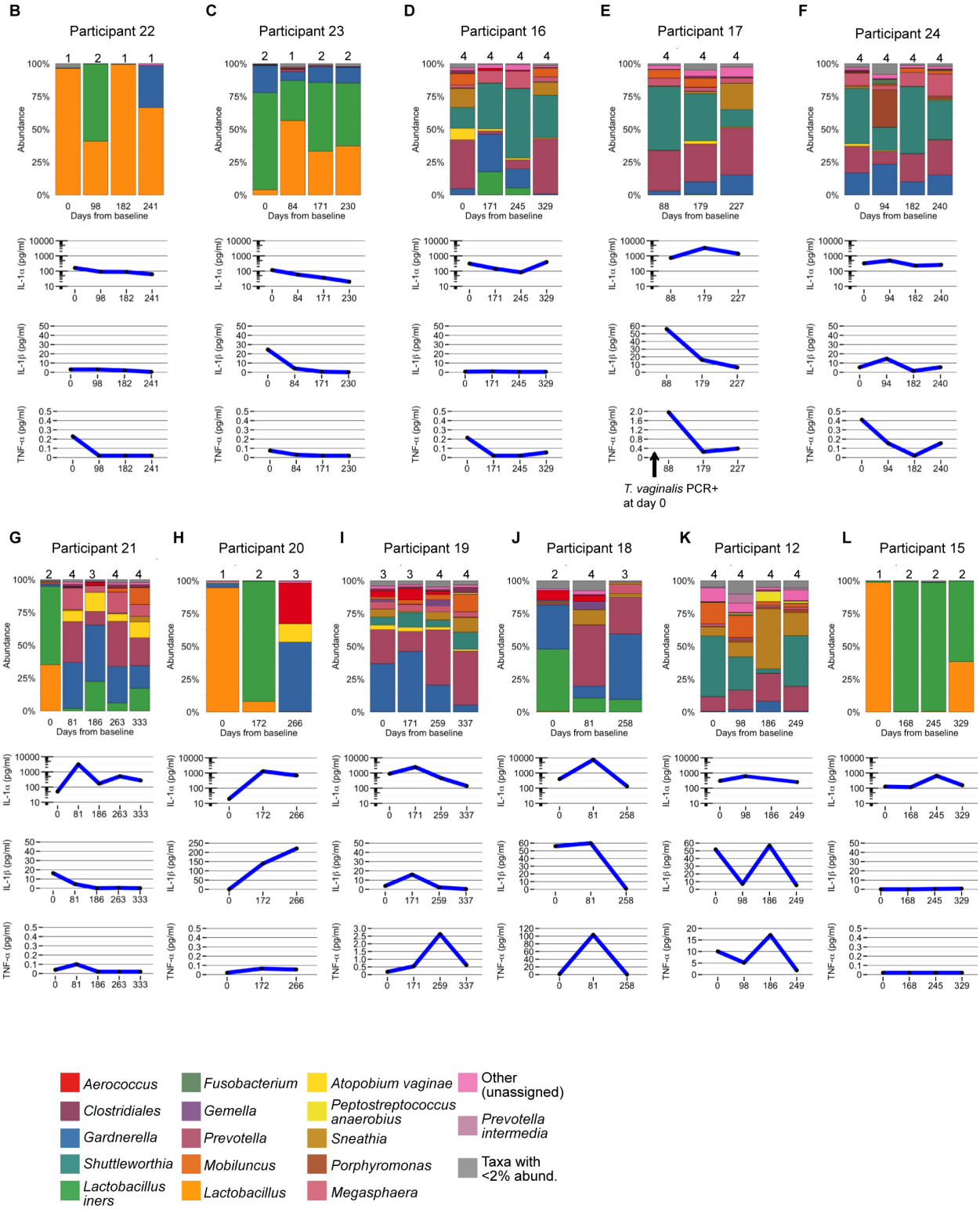


Figure S5, Related to Figure 5. Longitudinal bacterial abundances and genital cytokine levels.

(**A**) Table of CT transitions between serial time points.

(**B-L**) Longitudinal profiling of cervical bacterial abundances and IL-1α, IL-1β, and TNF-α concentrations in five women with fairly constant microbial communities (**B-F**) and six women with community fluctuations (**G-L**). Cervicotype (CT) designations are shown above each barplot.

A

Parameter	Spearman r	95% confidence interval	P value (two-tailed)
Prevotella	0.4443	0.2588 to 0.5982	< 0.0001
Mycoplasma	0.4781	0.2985 to 0.6251	< 0.0001
Other (unassigned)	0.5134	0.3404 to 0.6528	< 0.0001
Gemella	0.5234	0.3525 to 0.6606	< 0.0001
Fusobacterium	0.3882	0.1944 to 0.5528	0.0001
Leptotrichiaceae	0.3911	0.1977 to 0.5552	0.0001
Lactobacillus iners	-0.3585	-0.5284 to -0.1610	0.0004
Sneathia	0.3504	0.1520 to 0.5217	0.0006
Aerococcus	0.3445	0.1454 to 0.5168	0.0007
Clostridiales	0.3422	0.1429 to 0.5149	0.0008
Mobiluncus	0.3347	0.1345 to 0.5086	0.001
Lactobacillus	-0.2918	-0.4725 to -0.08760	0.0045
Arcanobacterium	0.2628	0.05625 to 0.4477	0.0109
Atopobium vaginae	0.2471	0.03951 to 0.4342	0.017
Veillonella	0.2279	0.01917 to 0.4175	0.028
Porphyromonas	0.1957	-0.01449 to 0.3893	0.0601
Peptostreptococcus anaerobius	0.1733	-0.03764 to 0.3695	0.0966
Lactobacillus reuteri	-0.1455	-0.3446 to 0.06609	0.164
Prevotella intermedia	0.143	-0.06868 to 0.3423	0.1715
Gardnerella	0.107	-0.1049 to 0.3097	0.3071
Prevotella melaninogenica	0.1022	-0.1097 to 0.3053	0.3296
Shuttleworthia	0.08394	-0.1279 to 0.2885	0.4238
Megasphaera	0.07931	-0.1325 to 0.2842	0.4498
Prevotella copri	-0.05216	-0.2589 to 0.1592	0.6195

B

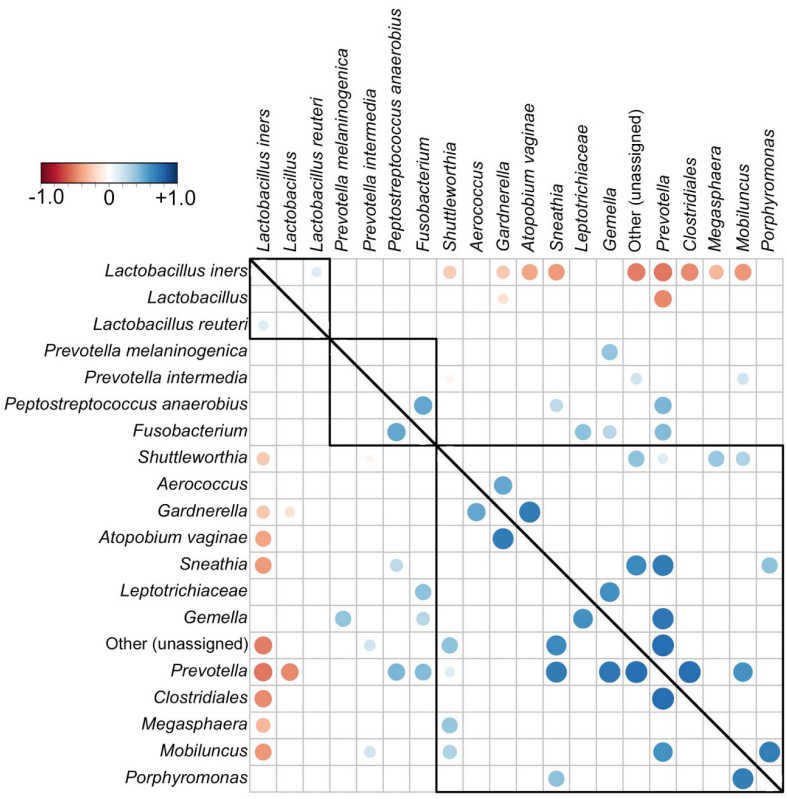


Figure S6, Related to Figure 6. Correlations of bacterial communities with genital cytokines, and bacterial species with one another.

(**A**) Spearman correlations between individual bacteria and cytokine PC1.

(**B**) Graphical Spearman correlation matrix of bacterial abundances found in the 94 cervical swabs. The areas and colors of circles show the absolute value of the corresponding correlation coefficients. Only correlations with $p < 0.01$ are shown.

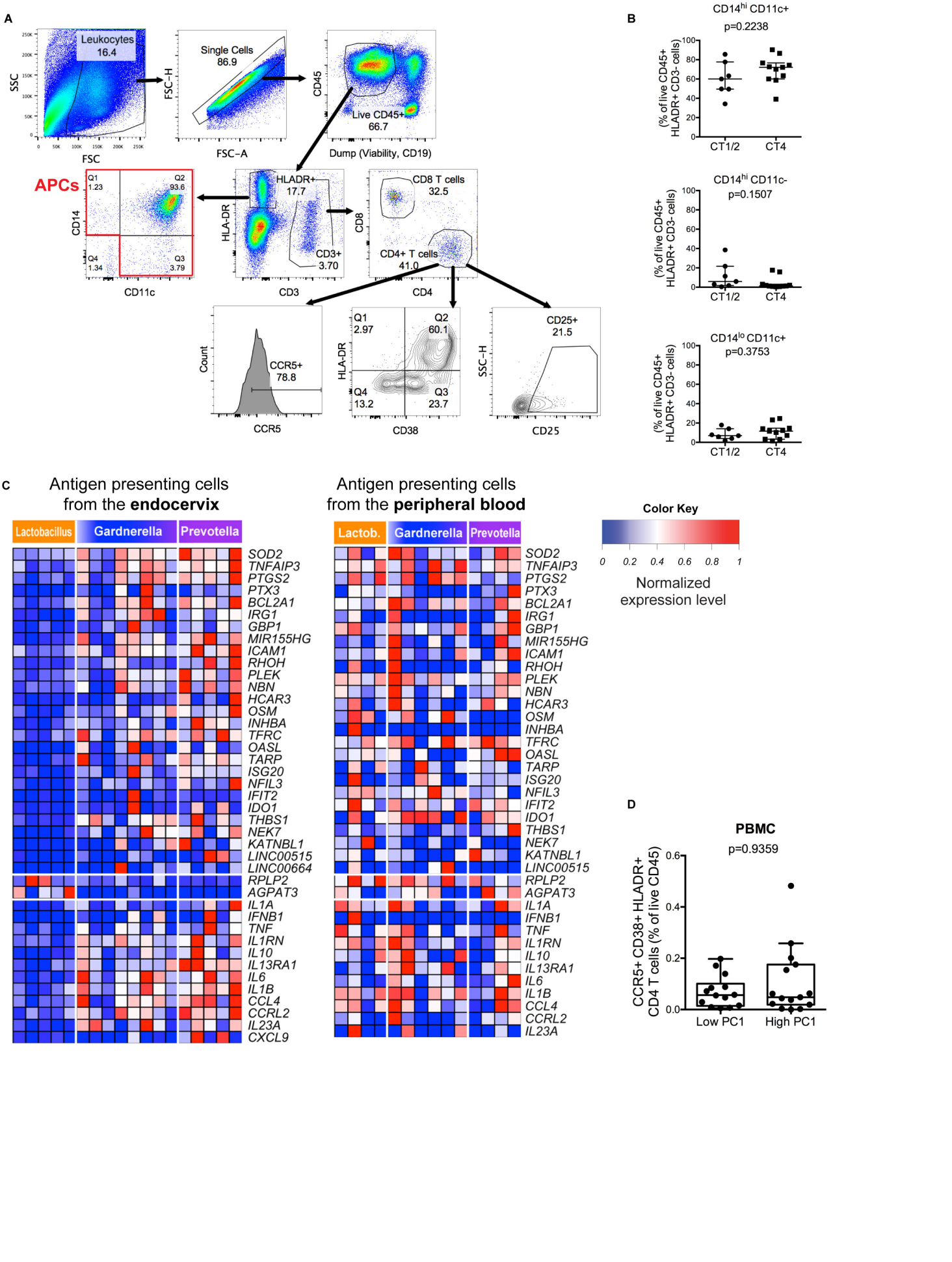

A
Leukocytes
16.4
SSC
FSC
Single Cells
86.9
FSC-H
FSC-A
CD45
Live CD45+
66.7
Dump (Viability, CD19)
APCs
Q1
1.23
Q2
93.6
Q4
1.34
Q3
3.79
CD14
CD11c
HLADR+
17.7
HLA-DR
CD3+
3.70
CD3
CD8 T cells
32.5
CD8
CD4+ T cells
41.0
CD4
Count
CCR5+
78.8
CCR5
Q1
2.97
Q2
60.1
Q4
13.2
Q3
23.7
HLA-DR
CD38
CD25+
21.5
SSC-H
CD25
B
CD14hi CD11c+
p=0.2238
(% of live CD45+ HLADR+ CD3- cells)
CT1/2
CT4
CD14hi CD11c-
p=0.1507
(% of live CD45+ HLADR+ CD3- cells)
CT1/2
CT4
CD14lo CD11c+
p=0.3753
(% of live CD45+ HLADR+ CD3- cells)
CT1/2
CT4
c
Antigen presenting cells from the endocervix
Lactobacillus
Gardnerella
Prevotella
Antigen presenting cells from the peripheral blood
Lactob.
Gardnerella
Prevotella
SOD2
TNFAIP3
PTGS2
PTX3
BCL2A1
IRG1
GBP1
MIR155HG
ICAM1
RHOH
PLEK
NBN
HCAR3
OSM
INHBA
TFRC
OASL
TARP
ISG20
NFIL3
IFIT2
IDO1
THBS1
NEK7
KATNBL1
LINC00515
LINC00664
RPLP2
AGPAT3
IL1A
IFNB1
TNF
IL1RN
IL10
IL13RA1
IL6
IL1B
CCL4
CCRL2
IL23A
CXCL9
Color Key
0
0.2
0.4
0.6
0.8
1
Normalized expression level
D
PBMC
p=0.9359
CCR5+ CD38+ HLADR+ CD4 T cells (% of live CD45)
Low PC1
High PC1

Figure S7, Related to Figure 7. No difference in peripheral blood antigen presenting cells from women with different genital bacterial communities.

(**A**) Gating strategy for flow cytometric analysis. Fluorescence minus one controls were used to determine the positive gates for CD38, HLA-DR, CCR5, and CD25.

(**B**) Comparing APC subset frequency between CT1/2 and CT4.

(**C**) Left: heatmap displaying normalized gene expression values of genes that were significantly differentially expressed in cervical antigen presenting cells when comparing CT1/2 vs. CT4 (**Figure 7C**). *Gardnerella* containing populations are shown for comparison. Right: Heatmap displaying normalized gene expression values from peripheral blood APCs of genes that were significantly differentially expressed in cervical antigen presenting cells when comparing CT1/2 vs. CT4 (**Figure 7C**).

(**D**) Number of live PBMC CD38+ HLADR+ CCR5+ CD4 T cells from women in the highest and lowest quintile of cytokine PC1 (Mann-Whitney), matching the women shown in **Figure 7D**.

Tables S1-S6:

Table S1, Related to Figure 1. Contraceptive usage, active sexually transmitted infections, condom use, and sexual behavior are not associated with the bacterial community state.

	All n=94	**CT1 n=8**	**CT2 n=30**	**CT3 n=25**	**CT4 n=31**	
STIs:	n (%)	n (%)	n (%)	n (%)	n (%)	
Gonorrhea	1 (1%)	1 (12%)	0 (0%)	0 (0%)	0 (0%)	p=0.08391 [a]
Chlamydia	9 (9.6%)	0 (0%)	4 (13%)	2 (8%)	3 (9.7%)	p=0.8855 [a]
Trichomonas	6 (6.4%)	0 (0%)	1 (3%)	3 (12%)	2 (6.5%)	p=0.5796 [a]
HSV2	3 (3.1%)	0 (0%)	0 (0%)	1 (4%)	2 (6.5%)	p=0.6197 [a]
HSV1	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	p=1.000 [a]
No STIs	71 (75%)	7 (88%)	25 (83%)	19 (76%)	20 (64%)	p=0.3609 [b]
History of STIs (self-reported):						
Yes	18 (19%)	1 (12%)	10 (33%)	4 (16%)	3 (9.7%)	p=0.09167 [c]
No	72 (77%)	7 (88%)	18 (60%)	19 (76%)	28 (90%)	
Not available	4 (4.2%)	0 (0%)	2 (6%)	2 (8%)	0 (0%)	
Experiencing STI symptoms at time of exam: [d]						
Yes	2 (2.1%)	0 (0%)	1 (3.3%)	0 (0%)	1 (3.2%)	p=1 [c]
No	89 (95%)	8 (100%)	28 (93.3%)	24 (96%)	29 (93.5%)	
Not available	3 (3.1%)	0 (0%)	1 (3.3%)	1 (4%)	1 (3.2%)	
Contraceptive:						
No family planning	43 (46%)	4 (50%)	14 (47%)	10 (40%)	15 (48%)	p=0.9474 [c]
DMPA/ Nuristerate	42 (45%)	4 (50%)	14 (47%)	12 (48%)	12 (38%)	
OCP/ Non-hormonal contra.	9 (9.6%)	0 (0%)	2 (6%)	3 (12%)	4 (13%)	
Condoms in 30 days:						
Always	16 (17%)	3 (37%)	3 (10%)	4 (16%)	6 (19%)	p=0.2775 [c]
Sometimes	29 (31%)	4 (50%)	10 (33%)	7 (28%)	8 (25%)	
Never	27 (29%)	0 (0%)	12 (40%)	6 (24%)	9 (29%)	
No response/ no sex in past 30 days	20 (21%)	1 (13%)	3 (10%)	8 (32%)	8 (26%)	
# of sex episodes in past 7 days:						
None	47 (50%)	4 (50%)	10 (33%)	15 (60%)	18 (58%)	p=0.2633 [c]
One or more	43 (46%)	4 (50%)	17 (57%)	9 (36%)	13 (42%)	
Not available	4 (4%)	0 (0%)	3 (10%)	1 (4%)	0 (0%)	
# of regular sex partners in last 30 days:						
None	39 (41%)	3 (38%)	7 (23%)	15 (62%)	14 (45%)	p=0.05377 [c]
One	51 (54%)	4 (50%)	21 (70%)	9 (38%)	17 (555)	
More than one	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	
Not available	4 (5%)	1 (12%)	2 (6.6%)	1 (4%)	0 (0%)	
# of casual sex partners in last 30 days:						
None	88 (94%)	7 (88%)	26 (93%)	24 (96%)	31 (100%)	p=0.3131 [c]
One	2 (21%)	0 (0%)	2 (6.6%)	0 (0%)	0 (0%)	
More than one	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	
Not available	4 (5%)	1 (12%)	2 (6.6%)	1 (4%)	0 (0%)	
Days since last sexual act (Median, IQR):	8 [3, 26]	8 [6, 21]	6 [2, 19.75]	16.5 [4, 37]	14 [5, 52]	p=0.275 [e]
# of sex episodes in past 30 day (Median, IQR):	2 [0, 4]	2 [1, 4]	2 [1,4]	1.5 [0, 3]	1 [0,4]	p=0.289 [e]
# of anal sex acts in past 30 days (Median, IQR):	0 [0, 0]	0 [0, 0]	0 [0, 0]	0 [0, 0]	0 [0, 0]	p=0.621 [e]
Age of participant (Median, IQR)	21 [20,22]	21 [19.25, 22]	21 [20.75, 23]	21 [20, 22]	21 [20, 23]	p=0.5631 [e]

a. Fisher's exact test, comparing each STI to those without that STI.
b. Fisher's exact test, comparing those with any STI to those without an STI.
c. Fisher's exact test, comparing all the groups lists under the subheading.
d. Symptoms defined as itching, pain, burning, sores, foul smelling discharge, blood, and pain during sex.
e. Kruskal-Wallis test, shown with median and interquartile range.

Table S2, Related to Figure 3. Demographics, sexual behavior, and progesterone levels in women in the top and bottom quartile of cytokine PC1. Women in the top quartile of cytokine PC1 had older current partners (1.91 year difference in means, $P = 0.0103$). Data shown as median with interquartile range or mean with standard deviation.

	Low PC1	High PC1	Significance
Age of sexual debut	18 [17, 19]	18 [17, 19]	p = 0.9879 [a]
Age of first sexual partner	21 [19, 22]	21 [20, 22.5]	p = 0.2986 [a]
Age of oldest sexual partner	26 [23, 27]	26 [24, 28]	p = 0.4199 [b]
Age of current sexual partner	23 [21, 26]	26 [23.5, 28]	p = 0.0103 [b]
Days since last sex episode	13 [4, 42]	12 [5.75, 42]	p = 0.9528 [b]
Number of sex acts in last 7 days	0 [0, 1]	0 [0,1]	p = 0.8735 [b]
Number of sex acts in last 30 days	2 [0, 4]	2 [0,4]	p = 0.8619 [b]
Blood progesterone level (ng/ml)	0.15 [0.1, 0.5]	0.1 [0.1, 0.2]	p = 0.4008 [b]

	Low PC1	High PC1	Significance [c]
No sex partners in last 7 days	16	12	p = 0.2167
1 or more sex partners	15	22	
No casual sex partners in last 7 days	31	34	p = 1
1 or more casual sex partners	0	0	
No oral sex	28	31	p = 1
Oral sex reported	2	2	
Always used condoms in past 30 days, or no sex	9	12	p = 0.553
Sometimes used condoms	12	11	
Never used condoms	6	11	
No family planning	19	15	p = 0.2162
Using family planning	12	19	
No STI history	27	23	p = 1
STI history	7	7	
No STI symptoms	30	32	p = 1
STI symptoms	1	2	

[a] Unpaired T test

[b] Mann-Whitney Test

[c] Fisher's exact test

Table S3, Related to Figure 5. Metadata from thirteen participants with longitudinal 16S profiling and cervicovaginal lavage measurements.

	Days from baseline	Cervico-type	STI PCR	BV	# Sex episodes in 7 days	Days since last sex episode	# anal sex episodes in last 30 days	Condom use in last 30 days	History of STIs (self-reported)	STI symptoms now (self-reported)	Family planning (FP) method
Participant 12	0	4	------	6	0	14	0	Always	No	No	IUD
Participant 12	98	4	------	n/a	0	16	0	Always	No	Yes	IUD
Participant 12	186	4	------	n/a	1	1	0	Always	No	No	IUD
Participant 12	249	4	NG+	7	1	2	0	Never	No	No	IUD
Participant 13	0	3	------	0	2	5	0	Sometimes	No	No	Depoprovera
Participant 13	87	2	------	1	2	7	0	Sometimes	No	No	Depoprovera
Participant 13	178	2	------	n/a	2	3	0	Sometimes	No	No	Depoprovera
Participant 13	269	2	------	1	n/a	n/a	n/a	n/a	n/a	n/a	Depoprovera
Participant 13	335	2	------	0	0	12	0	Sometimes	No	No	Depoprovera
Participant 14	0	2	------	0	1	7	0	Sometimes	No	No	No FP
Participant 14	81	2	n/a	n/a	3	4	0	Never	No	No	No FP
Participant 14	179	4	------	n/a	1	3	0	Never	No	No	No FP
Participant 14	227	4	NG+	1	1	6	0	Never	No	No	No FP
Participant 15	0	1	------	0	1	7	0	Never	No	No	Nuristerate
Participant 15	168	2	------	n/a	1	4	0	Never	No	No	Nuristerate
Participant 15	245	2	------	1	4	3	10	Never	No	No	Nuristerate
Participant 15	329	2	------	1	4	2	0	Sometimes	Yes	No	Nuristerate
Participant 16	0	4	------	9	0	60	0	Sometimes	No	No	No FP
Participant 16	171	4	------	n/a	0	53	0	n/a	No	No	No FP
Participant 16	245	4	------	6	0	26	0	Always	No	No	No FP
Participant 16	329	4	------	n/a	0	8	0	Never	No	No	No FP
Participant 17	0	n/a	TV+	3	1	6	0	Always	No	No	Oral contraceptives
Participant 17	88	4	n/a	n/a	1	7	0	Sometimes	No	No	Oral contraceptives
Participant 17	179	4	HSV2 +	9	1	4	0	Never	No	No	Oral contraceptives
Participant 17	227	4	n/a	n/a	5	2	0	Sometimes	No	No	Oral contraceptives
Participant 18	0	2	------	1	0	15	0	Sometimes	No	No	No FP
Participant 18	81	4	n/a	n/a	1	6	0	Sometimes	No	No	Oral contraceptives
Participant 18	258	3	CT+	1	0	45	n/a	n/a	No	No	Oral contraceptives
Participant 19	0	3	------	6	2	3	0	Sometimes	No	No	Nuristerate
Participant 19	171	3	------	n/a	0	17	1	Never	No	No	No FP
Participant 19	259	4	------	9	0	34	n/a	n/a	n/a	No	No FP
Participant 19	337	4	------	n/a	0	17	0	Sometimes	No	No	No FP
Participant 20	0	1	------	0	1	7	0	Sometimes	No	No	No FP
Participant 20	172	2	------	1	0	22	0	Sometimes	No	No	No FP
Participant 20	266	3	------	1	0	18	0	Sometimes	No	No	No FP
Participant 21	0	2	------	0	1	6	0	Sometimes	Yes	No	Depoprovera
Participant 21	81	4	------	3	1	4	0	Sometimes	No	No	No FP
Participant 21	186	3	------	n/a	1	4	0	Never	No	No	No FP
Participant 21	263	4	------	9	0	30	0	Never	No	No	No FP
Participant 21	333	4	NG+	6	0	16	0	Sometimes	No	No	No FP
Participant 22	0	1	------	1	1	6	0	Never	No	No	No FP
Participant 22	98	2	------	n/a	1	5	0	Never	No	No	No FP
Participant 22	182	1	------	n/a	3	1	0	Sometimes	No	No	No FP
Participant 22	241	1	NG+, CT+	2	1	3	0	Never	No	No	Nuristerate
Participant 23	0	2	------	2	0	30	0	Never	Yes	No	Depoprovera
Participant 23	84	1	n/a	n/a	0	60	0	n/a	No	No	Depoprovera
Participant 23	171	2	------	n/a	0	30	0	Sometimes	No	No	Depoprovera
Participant 23	230	2	------	1	0	60	n/a	n/a	No	No	Depoprovera
Participant 24	0	4	------	8	1	3	5	Never	No	No	No FP
Participant 24	94	4	------	n/a	0	27	0	Sometimes	No	No	No FP
Participant 24	182	4	------	n/a	0	9	0	Never	No	No	No FP
Participant 24	240	4	------	6	1	4	1	Sometimes	No	Yes	No FP

Table S4, Related to Figure 6. IL-1α, IL-1β, TNF-α, and IL-8 induced by *in vitro* application of bacterial species to vaginal epithelial cells. MALP-2 (25nM) is used as a positive control. Data shown as mean ± SEM (performed in duplicate). We were not able to test *Prevotella amnii* at a greater concentration than 2 log_{10}CFU, but it elicited significantly higher TNF-a and IL-8 secretion than 2 log_{10}CFU of *L. crispatus.*

IL-1α (pg/ml)

log_{10}CFU/cm^2	*Lactobacillus crispatus*	*Lactobacillus iners*	*Aerococcus christensenii*	*Mobiluncus mulieris*	*Prevotella amnii*	*Sneathia amnii*	*Sneathia sanguinegens*
15	53 ± 5.0	67 ± 13	5 ± 0	121.5 ± 17.5	n/a	292.5 ± 13.5	386.5 ± 3.50
12	51.5 ± 1.50	56.5 ± 2.50	2.5 ± 1.5	66 ± 4.0	n/a	263 ± 71.0	252.5 ± 12.5
8	48 ± 1.0	46 ± 6.0	6 ± 0	49.5 ± 5.50	n/a	87.5 ± 13.5	96.5 ± 2.50
5	56.5 ± 6.5	43.5 ± 2.5	0.5 ± 0.5	46.5 ± 4.50	n/a	48.5 ± 1.50	70.5 ± 10.5
2	52.5 ± 3.5	45 ± 3.0	18.5 ± 16.5	46.5 ± 1.50	83 ± 28	47 ± 2.0	43.5 ± 1.5

IL-1β (pg/ml)

log_{10}CFU/cm^2	*Lactobacillus crispatus*	*Lactobacillus iners*	*Aerococcus christensenii*	*Mobiluncus mulieris*	*Prevotella amnii*	*Sneathia amnii*	*Sneathia sanguinegens*
15	6.0 ± 3.0	21.5 ± 2.5	5.5 ± 2.5	18.5 ± 6.5	n/a	77.5 ± 17.5	135 ± 3.0
12	3.0 ± 0.0	8.0 ± 0.0	4.0 ± 1.0	5.0 ± 1.0	n/a	37 ± 0.0	65 ± 21.0
8	5.5 ± 0.5	4.0 ± 2.0	3.0 ± 1.0	3.5 ± 1.5	n/a	16.5 ± 3.5	20 ± 3.0
5	4.5 ± 0.5	3.0 ± 0.0	5.0 ± 0.0	3.5 ± 0.5	n/a	4.0 ± 1.0	4.5 ± 0.5
2	4.0 ± 0.0	2.5 ± 0.5	3.0 ± 1.0	2.0 ± 0.0	5.5 ± 1.5	3.0 ± 0.0	3.5 ± 1.5

TNF-α (pg/ml)

log_{10}CFU/cm^2	*Lactobacillus crispatus*	*Lactobacillus iners*	*Aerococcus christensenii*	*Mobiluncus mulieris*	*Prevotella amnii*	*Sneathia amnii*	*Sneathia sanguinegens*
15	4.5 ± 0.5	25.5 ± 2.5	5.0 ± 0.0	263.5 ± 184.5	n/a	76 ± 40	26 ± 14
12	2.5 ± 0.5	11 ± 0	2.5 ± 1.5	53.5 ± 5.50	n/a	156.5 ± 137.5	107 ± 59.0
8	0 ± 0.5	9.0 ± 1.0	6.0 ± 0.0	28.5 ± 22.5	n/a	10 ± 4.0	20 ± 2.0
5	0.5 ± 0.5	4.0 ± 1.0	0.5 ± 0.5	5.5 ± 1.5	n/a	0 ± 0	18.5 ± 6.5
2	0.5 ± 0.5	19 ± 12	18.5 ± 16.5	3 ± 2.0	69 ± 11	2.5 ± 0.5	4.0 ± 2.0

IL-8 (pg/ml)

log_{10}CFU/cm^2	*Lactobacillus crispatus*	*Lactobacillus iners*	*Aerococcus christensenii*	*Mobiluncus mulieris*	*Prevotella amnii*	*Sneathia amnii*	*Sneathia sanguinegens*
15	581 ± 1.00	4,173 ± 167	645 ± 8.00	10,100 ± 875	n/a	13,800 ± 744	11,300 ± 578
12	363.5 ± 2.50	1,690 ± 10.0	364 ± 12.0	2,580 ± 93.5	n/a	8,560 ± 505	10,900 ± 205
8	339.5 ± 0.500	568 ± 7.00	314.5 ± 4.50	481 ± 43.5	n/a	2,060 ± 47.0	6,490 ± 121
5	360.5 ± 3.50	303 ± 1.00	315.5 ± 3.50	302 ± 17.0	n/a	403 ± 19	837 ± 58.5
2	344.5 ± 28.5	301 ± 13.0	312.5 ± 14.5	289 ± 5.00	1,290 ± 82.5	275 ± 4.0	305 ± 19.0

Table S5, Related to Figure 7. Bacterial pathways identified by HUMAnN to be differentially abundant between women with high (n=6; 1 in CT3, 5 in CT4) versus low (n=6; 2 in CT1, 3 in CT2, 1 in CT3) genital inflammation. The bacterial communities of the same 12 participants are also represented in **Figure S2**.

KEGG Pathway	Enriched in:	LDA Score (Log 10)
ko00780| Biotin metabolism	HighInflamm	3.290
ko00540| Lipopolysaccharide biosynthesis	**HighInflamm**	**3.201**
ko00750| Vitamin B6 metabolism	HighInflamm	3.157
ko00410| beta-Alanine metabolism	HighInflamm	3.149
ko02040| Flagellar assembly	HighInflamm	3.110
ko02030| Bacterial chemotaxis	HighInflamm	3.088
ko00523| Polyketide sugar unit biosynthesis	HighInflamm	3.042
ko05310| Asthma	HighInflamm	3.042
ko00660| C5-Branched dibasic acid metabolism	HighInflamm	2.986
ko00860| Porphyrin and chlorophyll metabolism	HighInflamm	2.915
ko00340| Histidine metabolism	HighInflamm	2.885
ko00260| Glycine, serine and threonine metabolism	HighInflamm	2.871
ko00361| Chlorocyclohexane and chlorobenzene degradation	HighInflamm	2.850
ko00330| Arginine and proline metabolism	HighInflamm	2.715
ko05131| Shigellosis	HighInflamm	2.525
ko00562| Inositol phosphate metabolism	HighInflamm	2.490
ko05120| Epithelial cell signaling in Helicobacter pylori infection	HighInflamm	2.311
ko02060| Phosphotransferase system (PTS)	LowInflammation	3.503
ko03430| Mismatch repair	LowInflammation	3.227
ko00240| Pyrimidine metabolism	LowInflammation	3.221
ko00051| Fructose and mannose metabolism	LowInflammation	3.215
ko00072| Synthesis and degradation of ketone bodies	LowInflammation	3.112
ko00052| Galactose metabolism	LowInflammation	3.100
ko00520| Amino sugar and nucleotide sugar metabolism	LowInflammation	3.060
ko00230| Purine metabolism	LowInflammation	2.962
ko00010| Glycolysis / Gluconeogenesis	LowInflammation	2.903
ko00561| Glycerolipid metabolism	LowInflammation	2.865
ko03020| RNA polymerase	LowInflammation	2.794
ko00621| Dioxin degradation	LowInflammation	2.728
ko05150| Staphylococcus aureus infection	LowInflammation	2.640
ko02010| ABC transporters	LowInflammation	2.603
ko04662| B cell receptor signaling pathway	LowInflammation	2.331

Supplemental Experiment Procedures:

Study cohort eligibility criteria

To be eligible for the study, participants had to be female, 18-23 years old, HIV-negative, able to understand the information and consent forms, willing to adhere to study requirements, willing to have HIV testing performed twice-weekly, and willing to have samples stored. Participants could not be pregnant, anemic, or enrolled in any other study.

Sample selection

At the time of analysis, samples were available from 146 study participants who had completed at least one mucosal sampling, and 13 participants who had completed at least three samplings. Of these 146 participants, we performed 16S sequencing on all subjects who had a complete set of cervical swab, CVL, and cellular phenotyping data (n=94).

Sample collection

A single nurse performed all the pelvic exams. Ectocervical and midvaginal swabs were immediately placed into sterile cryovials after collection and stored at 4°C for 1-3 hours during transport to the laboratory, followed by long-term storage at -80°C. The first posterior fornix swab was sent to Global Labs (Durban, South Africa) for sexually transmitted infection testing. At the first and fourth pelvic exam visits, the second posterior fornix swab was rolled onto a glass slide, allowed to air dry, and a gram stain and Nugent scoring (Nugent et al., 1991) were performed by Global Labs.

The cervicovaginal lavage was performed by washing the cervicovaginal walls with 5 mL of sterile saline using a sterile Pasteur pipette. The lavage was kept at 4°C for 1-3 hours during transport to the laboratory, where it was centrifuged at 1700 rpm at 4°C to fractionate the cellular component from the supernatant. The supernatant was aliquoted and stored at -80°C, and subsequently used for cytokine measurements.

Cellular collection was performed by fully inserting a cytobrush (Cytobrush Plus GT, Cooper Surgical) into the endocervical canal and rotating twice exactly 360°. Immediately after collection, the cytobrush was placed into antibiotic supplemented RPMI media with 10% FCS and stored at 4°C for 1-3 hours. In the laboratory, the cells were dislodged from the cytobrush and washed prior to cell surface staining and FACS analysis.

Peripheral blood mononuclear cells were isolated from blood using a standard Ficoll gradient.

Sexually transmitted infection detection

Neisseria gonorrhoea and *Chlamydia trachomatis* were tested using the GeneXpert CT/NG assay (Cepheid). *Trichomonas vaginalis*, HSV-1, and HSV-2 were tested by real-time PCR on a LightCycler (Roche). All positive results for *Trichomonas,* HSV-1, and HSV-2 were retested, and inconsistent results were sent to an outside lab for a third confirmatory test. Genital HSV-1 results were negative in all women whose samples were used in this study and thus are not referred to in the manuscript. Women with positive STI results were referred for treatment at an outside facility.

Progesterone measurement

Plasma progesterone levels were measured by the Massachusetts General Hospital Clinical Laboratory Research Core (MGH CLR) using a Chemiluminescent Microparticle Immunoassay (CMIA) (Abbott Laboratories).

Nucleic acid extraction from cervical swabs

We chose to utilize ectocervical swabs for microbial profiling, though we also had vaginal swabs available, because of the closer proximity of the ectocervix to the endocervical canal where the cellular sampling was occurring. The ectocervix is also more relevant to HIV transmission, which is thought to be more likely to occur through the cervical epithelium than the multi-layered vaginal

epithelium (Haase, 2010). Regardless, cervical and vaginal microbial populations are very similar based on data from other studies (Smith et al., 2014) and our cohort (data not shown).

Swabs were thawed on ice, placed into a solution of phenol:chloroform:isoamyl alcohol (25:24:1, pH 7.9, Ambion), 20% sodium dodecyl sulfate (Fisher), Tris-EDTA buffer, and 0.1mm glass beads (BioSpec), and then vigorously rubbed against the walls of the tube. The solution was homogenized using a bead beater for 2 minutes. Following centrifugation at 6,000x*g* for 3 minutes, the aqueous phase was transferred to a clean tube and an equal volume of phenol:chloroform:isoamyl alcohol was added and mixed by vortexing. Following a second centrifugation at 14,000x*g* for 5 minutes, the aqueous phase was transferred to a clean tube. Nucleic acid was precipitated with isopropanol and 3M sodium acetate pH 5.5 (Ambion) at -20°C overnight and then centrifuged at maximum speed for 30 minutes at 4°C. The supernatant was discarded and the pellet was washed by adding 0.5mL of 100% ethanol, centrifuging for 15 minutes at 4°C, and discarding the supernatant. The nucleic acid pellet was allowed to dry and was resuspended in 20 µL of molecular grade Tris-EDTA buffer (Promega). The nucleic acid concentration was measured using a Nanodrop (Thermo Scientific).

PCR amplification and sequencing of the V4 region of bacterial 16S rRNA gene

We chose to sequence the V4 region of the 16S *rRNA* gene due to the availability of extensively validated barcoded primers that are optimized for Illumina sequencing. While the V4 region has tremendous discriminatory power between most genera, it does not distinguish between all *Lactobacillus* species at a 97% operational taxonomic unit (OTU) level due to sequence similarity in that region.

Each cervical nucleic acid sample was diluted to 100 ng/µL using ultra-pure water (MoBio) prior to PCR amplification. PCR amplification was performed using the Phusion High-fidelity polymerase (New England Biolabs) in triplicate, with each well containing 2 µL template DNA. A no-template control was performed for each barcode pair on the same plate. PCR reactions were

performed with the following settings: denaturation for 30 s at 98°C, 30 cycles of 98°C for 10 s, 57°C for 30 s, and 72°C for 12 s, with a final extension for 10 m at 72°C. Triplicate PCR reactions for each sample were then combined into a single volume. Amplicons for each sample and a no-template control with the same barcode pair were run on a 1.5% agarose gel to confirm the presence of a single band at roughly 300-350 bp in the amplicon and absence of a band in the no-template control. Amplicons were quantified by Picogreen (Invitrogen) and equal amounts from each sample were pooled. The amplicon pool was cleaned using an UltraClean PCR Clean-up Kit (MoBio). The library was sequenced with a single 1x300 bp run on an Illumina MiSeq at the Center for Genome Sciences at Washington University in St. Louis using a 10% PhiX spike-in and custom sequencing primers, as described previously (Caporaso et al., 2011).

16S sequence quality control and taxonomic classification

Of the 13,964,311 input sequences, 5,975,498 reads passed the quality filtering step (minimum quality score of 30) and were assigned to samples using split_libraries.py. Operational taxonomic units (OTUs) were assigned using open-reference picking (97% identity, Greengenes v.13.8) with default parameters except for 0.1% subsampling. The median sequence count/sample was 40,657 with a standard deviation of 14,318. With a rarefaction cut-off of 10,000, one sample containing 1,521 counts was excluded from further analysis. Taxa with fewer than 5 reads were filtered, and the remaining taxa were summarized at the species level.

The principal coordinates plot was generated in EMPeror using the weighted UniFrac (Lozupone and Knight, 2005; Vazquez-Baeza et al., 2013) method with rarefaction of 2,000 reads per sample. Alpha diversity was calculated using Faith's phylogenetic diversity using the PD_whole_tree QIIME command.

Cervicotype determination

"Cervicotypes" were defined as follows: CT1 has a higher percentage of reads assigned by QIIME open-reference OTU picking to the genera *Lactobacillus* (but not *L. iners*) than *L. iners*, *Gardnerella*, or *Prevotella*; CT2 has a higher percentage of reads from *L. iners* than *Lactobacillus*, *Gardnerella*, or *Prevotella*; CT3 has a higher percentage of reads from *Gardnerella* than *Lactobacillus*, *L. iners*, or *Prevotella*; and CT4 is the remaining samples that lack dominance in *Lactobacillus*, *L. iners,* and *Gardnerella*. Shotgun DNA sequencing and oligotyping analyses demonstrated that CT1 was mostly composed of *Lactobacillus crispatus.*

Metagenomic shotgun DNA library specifications and quality control

Shotgun DNA libraries were made from total nucleic acid isolated from 12 cervical swab samples; the same total nucleic acid was used for 16S amplification. The samples were chosen based on CVL cytokine levels (six with very high inflammation, six with very low inflammation). DNA was isolated from total nucleic acid using an AllPrep DNA/RNA Micro Kit (Qiagen). Methylated DNA was removed using a Microbiome DNA Enrichment kit (New England Biolabs) to enrich for microbial DNA. Microbial DNA was sheared using an S2 sonicator (Covaris) and DNA was sized and quantified using a high-sensitivity DNA Bioanalyzer chip (Agilent). DNA libraries were prepared using a NEBNext DNA Library Construction kit (New England Biolabs), following manufacturer's directions with the double-sided SPRI bead cleanup option. The final library was sequenced on a paired-end 250 bp Illumina MiSeq run. The final library had a median size of 480 bp and was sequenced on a paired-end 250 bp run using the Illumina MiSeq at the Center for Genome Sciences at Washington University in St. Louis. After demultiplexing, each sample had 703,954 - 1,148,817 reads. Quality control filtering and adapter trimming were performed using fastq-mcf (Aronesty, 2013), with a quality threshold of 25, 0.01% "N" percentage causing cycle removal, and minimum remaining sequence length of 25. The forward and reverse reads of each sample were concatenated prior to MetaPhlAn2 analysis.

Using MetaPhlAn2, we detected the same genera using 16S and WGS methods for every bacteria with an abundance over 1%, with the exception of *Sneathia* and *Shuttleworthia*, which are under-represented in the available marker database due to a paucity of full-genome reference sequences.

Oligotyping

Oligotyping analysis (Eren et al., 2011) was performed on 16S V4 sequencing reads that were assigned by QIIME to the *Lactobacillus* genera but not to a specific species. A total of 233,821 sequences were analyzed; all had Phred quality scores above 30 and were trimmed to 270 bases. Seven base locations of interest (57, 59, 115, 124, 155, 214, and 245) were used to determine oligotypes. The resulting oligotypes had maximum Shannon entropy scores less than 0.2, which is within sequencing noise, indicating that the majority of subtle nucleotide variation was captured. Representative oligotypes were searched in the nucleotide BLAST database; the two most abundant oligotype sequences were:

>Oligotype GCTGCAC
TACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGAAGAATA
AGTCTGATGTGAAAGCCCTCGGCTTAACCGAGGAACTGCATCGGAAACTGTTTTTCTTGAGTG
CAGAAGAGGAGAGTGGAACTCCATGTGTAGCGGTGGAATGCGTAGATATATGGAAGAACACCA
GTGGCGAAGGCGGCTCTCTGGTCTGCAACTGACGCTGAGGCTCGAAAGCATGGGTAGCGAAC
AGGATTAGAAACCCTAGTAG

>Oligotype GCTGCAT
TACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGAAGAATA
AGTCTGATGTGAAAGCCCTCGGCTTAACCGAGGAACTGCATCGGAAACTGTTTTTCTTGAGTG
CAGAAGAGGAGAGTGGAACTCCATGTGTAGCGGTGGAATGCGTAGATATATGGAAGAACACCA
GTGGCGAAGGCGGCTCTCTGGTCTGCAACTGACGCTGAGGCTCGAAAGCATGGGTAGTGAAC
AGGATTAGAAACCCTAGTAG

Cytokine measurements in CVLs

CVLs were thawed on ice, centrifuged for 10 min at 1,000x*g* and 4°C, and the undiluted supernatant was immediately assayed. IL-1β, IL-2, IL-4, IL-8, IL-10, IL-12p70, IFN-γ, and TNF-α were measured with a high-sensitivity kit. IL-1α, IL-1RA, IP-10, MIP-1α, MIP-1β, RANTES, sCD40L,

FTL-3L, and IFN-α2 were measured with a regular sensitivity kit. The multiplexed bead assay was performed according to the manufacturer's protocol. A quality control was included with every plate to ensure reproducibility. Cytokine measurements below the limit of detection were assigned to a value of half of the minimum detectable concentration for that cytokine. When generating the heatmap and correlation matrix, measurements for each cytokine were first normalized by taking the log10 of the cytokine concentration, subtracting the mean measurement for that cytokine, and dividing by the standard deviation. The cytokine heatmap was generated using the complete clustering method on the Euclidean distance matrix.

Bacterial strains and colonization assays

The *Lactobacillus crispatus* isolate was originally collected from a vaginal swab from a healthy woman (Onderdonk et al., 1987). *Lactobacillus iners* (28746), *Aerococcus christensenii* (28831), *Peptostreptococcus* (44095), *Prevotella amnii* (53648), *Sneathia amnii* (64370), and *Sneathia sanguinegens* (41628) were acquired from the Culture Collection University of Göteborg (Sweden). *Mobiluncus mulieris* (35243) was acquired from the American Type Culture Collection (ATCC, Manassas, VA). Bacteria were cultured on Brucella agar under anaerobic conditions and allowed to colonize human vaginal epithelial cells at 2 to 15 log10 CFU/cm^2 for 24h under anaerobic conditions as previously described (Fichorova et al., 2011). IL-1α, IL-1β, IL-8, and TNF-α were measured in 24h supernatants using a Meso Scale Discovery multiplex electrochemiluminescence assay (Fichorova et al., 2011). MALP-2 (macrophage-activating lipopeptide-2, a TLR 2/4 agonist; 25nM) was used as a positive control. *Atopobium vaginae*, *Gardnerella vaginalis*, *Prevotella bivia*, and *Lactobacillus gasseri* were not tested due to our prior results indicating that all induce significant inflammatory responses except *L. gasseri* (previously referred to as *L. acidophilus*) (Fichorova et al., 2013).

Immunophenotyping of cervical cytobrush samples

Monoclonal antibodies to the following human proteins were used in multiparameter FACS analysis: CD45 (HI30), CD3 (UCHT1), CD8 (SK1), HLA-DR (G46-6), CD38 (HIT2), CD25 (2A3), CCR5 (2D7), CD11c (B-ly6), CD14 (M5E2), CD19 (HIB19) from BD Biosciences; and CD4 (S3.5) from Life Technologies. All antibodies were titrated before use. Viable cells were identified by a LIVE/DEAD violet dead cell stain (Life Technologies). Cells were passed through a 70μm filter prior to analysis on the FACS Aria III (BD Biosciences). Rainbow beads (Spherotech) were used to ensure comparable fluorescence measurements between experiments. There were 84 women for whom we had both cervical cellular phenotyping information and CVL cytokine measurements.

Immunohistochemistry

Excess surgical tissue deemed histologically normal by a certified pathologist was acquired through Massachusetts General Hospital (IRB protocol 2010P000632). CD14 (1:200, Clone 7, Abcam) staining was performed on formalin-fixed, paraffin-embedded endocervical tissue sections after antigen retrieval in 1X DIVA Decloaker (Biocare Medical) for 2 minutes. CD11c (1:50, Clone B-ly6, BD) staining was performed on acetone-fixed frozen endocervical tissue sections. After primary antibody incubation, a secondary HRP-labeled goat anti-mouse was applied (EnVision+ System-HRP, Dako). The stain was developed with DAB and enhancement (Dako), followed by a hematoxylin counterstain.

RNA extraction from sorted cells

Briefly, Gen-Elute linear polyacrylamide (Sigma) was added to the RNA in TRIzol, followed by the addition of 0.2 volumes of chloroform and vortexing. After centrifuging at 14,000x*g* for 5 minutes, the aqueous phase was transferred to a clean tube. The remaining steps were performed as described above (in “Nucleic acid extraction from cervical swabs”), with final nucleic acid reconstitution in 5μL of nuclease-free water.

Transcriptional analysis with SCRB-seq

Briefly, Poly(A)+ mRNA from extracted total RNA were converted to cDNA decorated with universal adapters, sample-specific barcodes and unique molecular identifiers (UMIs) using a template-switching reverse transcriptase. Decorated cDNA from multiple samples were then pooled, amplified and prepared for multiplexed sequencing using a modified transposon-based fragmentation approach that enriched for 3' ends and preserved strand information.

All second sequence reads were aligned to a reference database consisting of all human RefSeq mRNA sequences (obtained from the UCSC Genome Browser hg19) and the human hg19 mitochondrial reference sequence using bwa version 0.7.4 with non-default parameter "-l 24". Digital gene expression (DGE) profiles were then generated by counting, for each microplate well and RefSeq gene, the number of unique UMIs associated with that gene in that well. Python scripts implementing the alignment and DGE derivation are available from the authors upon request.

When analyzing the digital gene expression values using EdgeR, counts were normalized by weighted trimmed mean of M-values and significance was tested with a general linearized model.

Supplemental References

Aronesty, E. (2013). Comparison of Sequencing Utility Programs. The Open Bioinformatics Journal *7*, 1-8.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N., and Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences of the United States of America *108 Suppl 1*, 4516-4522.

Eren, A.M., Zozaya, M., Taylor, C.M., Dowd, S.E., Martin, D.H., and Ferris, M.J. (2011). Exploring the diversity of Gardnerella vaginalis in the genitourinary tract microbiota of monogamous couples through subtle nucleotide variation. PloS one *6*, e26732.

Fichorova, R.N., Buck, O.R., Yamamoto, H.S., Fashemi, T., Dawood, H.Y., Fashemi, B., Hayes, G.R., Beach, D.H., Takagi, Y., Delaney, M.L.*, et al.* (2013). The Villain Team-Up or how Trichomonas vaginalis and bacterial vaginosis alter innate immunity in concert. Sexually transmitted infections *89*, 460-466.

Fichorova, R.N., Yamamoto, H.S., Delaney, M.L., Onderdonk, A.B., and Doncel, G.F. (2011). Novel vaginal microflora colonization model providing new insight into microbicide mechanism of action. mBio *2*, e00168-00111.

Haase, A.T. (2010). Targeting early infection to prevent HIV-1 mucosal transmission. Nature *464*, 217-223.

Lozupone, C., and Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. Applied and environmental microbiology *71*, 8228-8235.

Nugent, R.P., Krohn, M.A., and Hillier, S.L. (1991). Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. Journal of clinical microbiology *29*, 297-301.

Onderdonk, A.B., Zamarchi, G.R., Rodriguez, M.L., Hirsch, M.L., Munoz, A., and Kass, E.H. (1987). Qualitative assessment of vaginal microflora during use of tampons of various compositions. Applied and environmental microbiology *53*, 2779-2784.

Smith, W.L., Hedges, S.R., Mordechai, E., Adelson, M.E., Trama, J.P., Gygax, S.E., Kaunitz, A.M., and Hilbert, D.W. (2014). Cervical and vaginal flora specimens are highly concordant with respect to bacterial vaginosis-associated organisms and commensal lactobacillus species in women of reproductive age. Journal of clinical microbiology *52*, 3078-3081.

Vazquez-Baeza, Y., Pirrung, M., Gonzalez, A., and Knight, R. (2013). EMPeror: a tool for visualizing high-throughput microbial community data. GigaScience *2*, 16.