

Supplementary Materials

Metagenomic Insight into the Vaginal Microbiome in Women Infected with HPV 16 and 18

Da-Ryung Jung¹⁺, Ye Seul Choi^{2,3}, Minsoo Jeong¹, Vineet Singh¹, Se Young Jeon^{4,5}, Incheol Seo^{5,6}, Nora Jee-Young Park^{5,7}, Yoon Hee Lee⁴, Ji Young Park⁷, Hyung Soo Han^{2,3,5,8}, Jae-Ho Shin^{1,9*} & Gun Oh Chong^{4,5*}

¹Department of Applied Biosciences, Kyungpook National University, Daegu 41566, Republic of Korea

²Department of Biomedical Science, Graduate School, Kyungpook National University, Daegu 41944, Republic of Korea

³BK21 Four Program, School of Medicine, Kyungpook National University, Daegu 41944, Republic of Korea

⁴Department of Obstetrics and Gynecology, Kyungpook National University Chilgok Hospital, Daegu 41404, Republic of Korea

⁵Clinical Omics Institute, Kyungpook National University, Daegu 41405, Republic of Korea

⁶Department of Immunology, School of Medicine, Kyungpook National University, Daegu 41944, Republic of Korea

⁷Department of Pathology, Kyungpook National University Chilgok Hospital, Daegu 41404, Republic of Korea

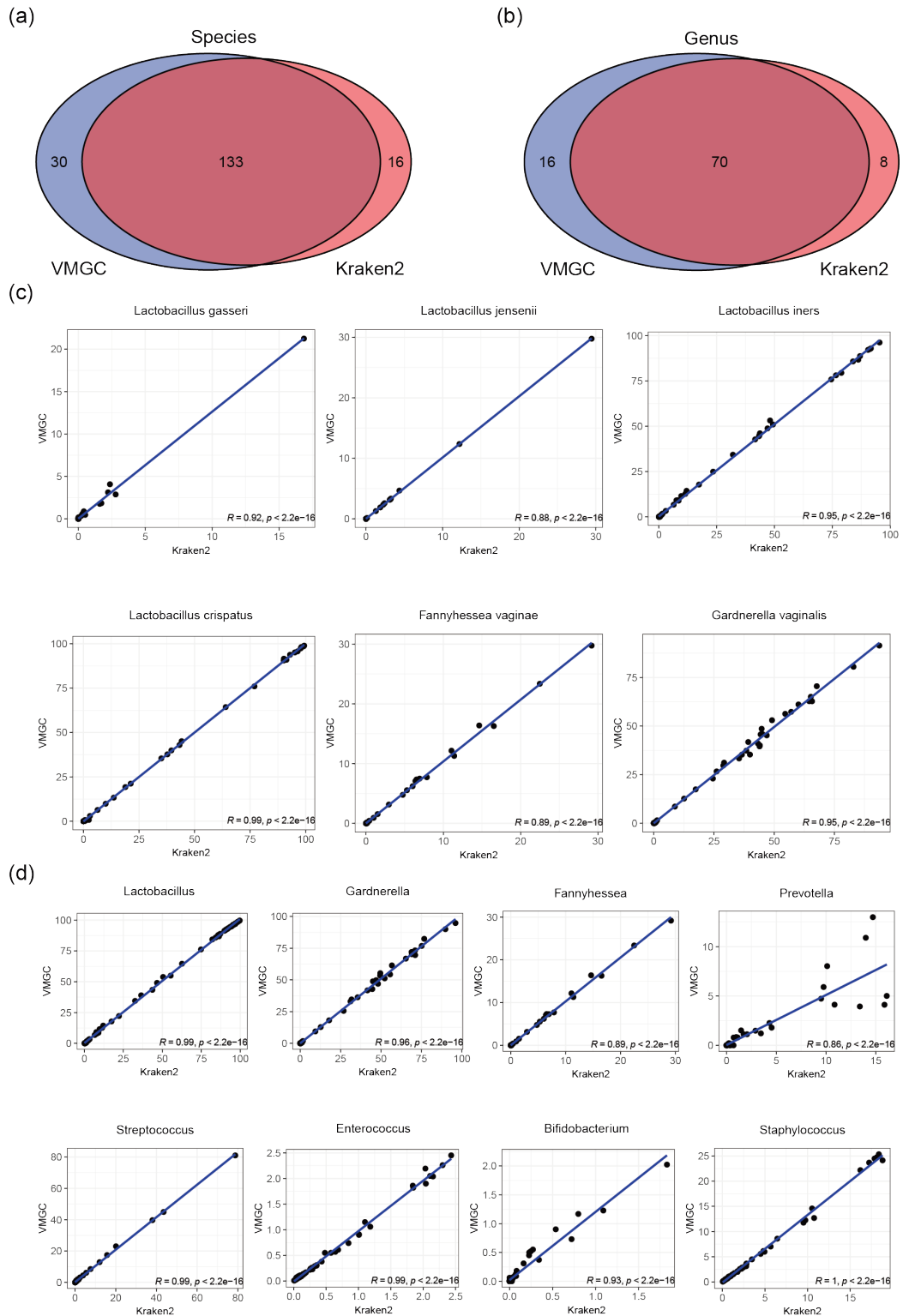
⁸Department of Physiology, School of Medicine, Kyungpook National University, Daegu 41944, Republic of Korea

⁹NGS Core Facility, Kyungpook National University, Daegu 41566, Republic of Korea

Authors' email addresses

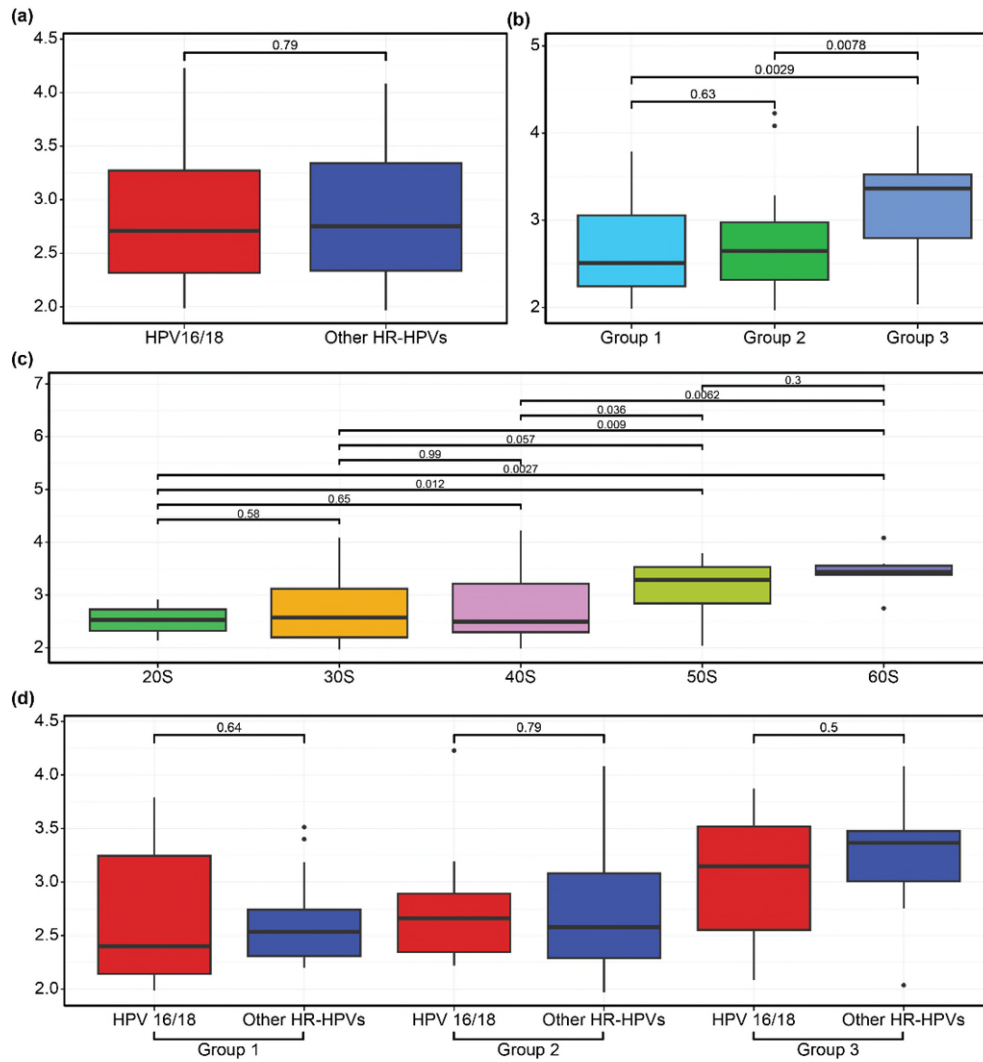
Da-Ryung Jung (D.R.J.), amugae1210@knu.ac.kr; Ye Seul Choi (Y.S.C.), yeseul.choi830@knu.ac.kr; Minsoo Jeong (M.J.), minsoo0326@knu.ac.kr; Vineet Singh (V.S.), singhvineet@knu.ac.kr; Se Young Jeon (S.Y.J.), tpdud0927@knu.ac.kr; Incheol Seo (I.S.), iseo@knu.ac.kr; Nora Jee-Young Park (N.J.Y.P.), pathpjy@knu.ac.kr; Yoon Hee Lee (Y. H. L.), yhlee1017@knu.ac.kr; Ji Young Park (J. Y. P.), jyparkmd@knu.ac.kr; Hyung Soo Han (H.S.H.), hshan@knu.ac.kr; Jae-Ho Shin (J.H.S.), jhshin@knu.ac.kr; Gun Oh Chong, gochong@knu.ac.kr (G.O.C.)

*Correspondence: Jae-Ho Shin, jhshin@knu.ac.kr; Gun Oh Chong, gochong@knu.ac.kr

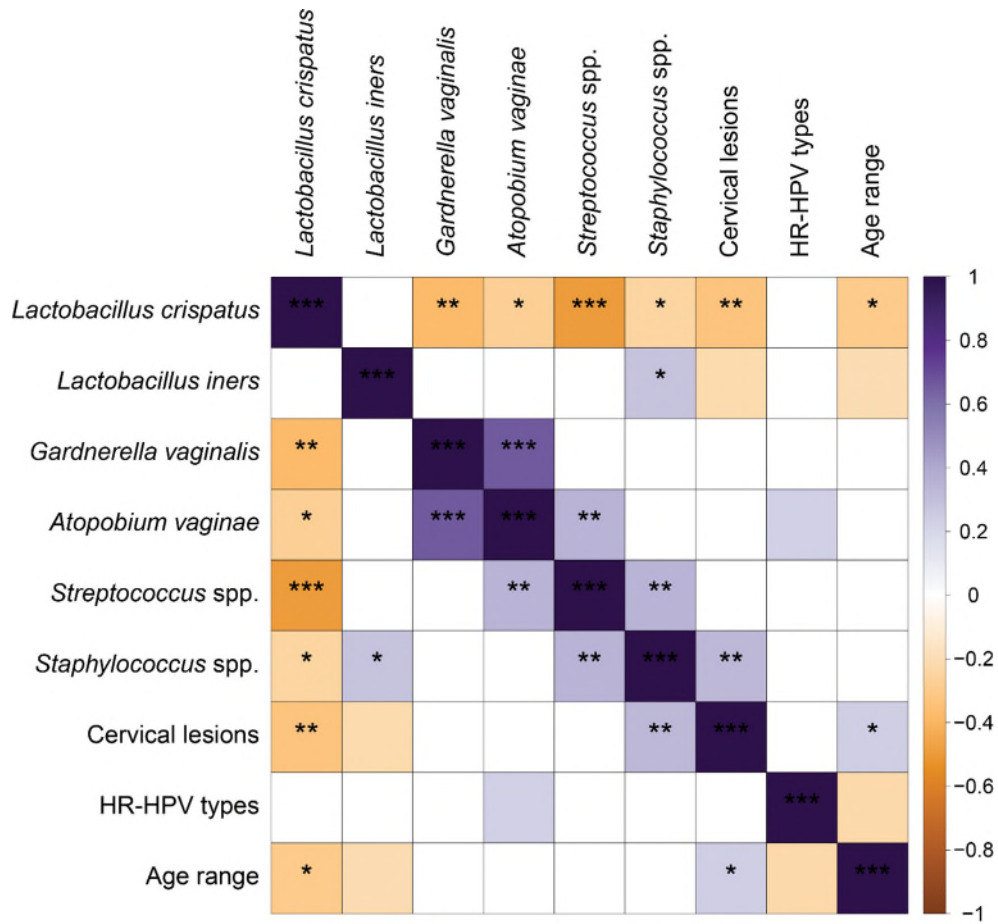


30

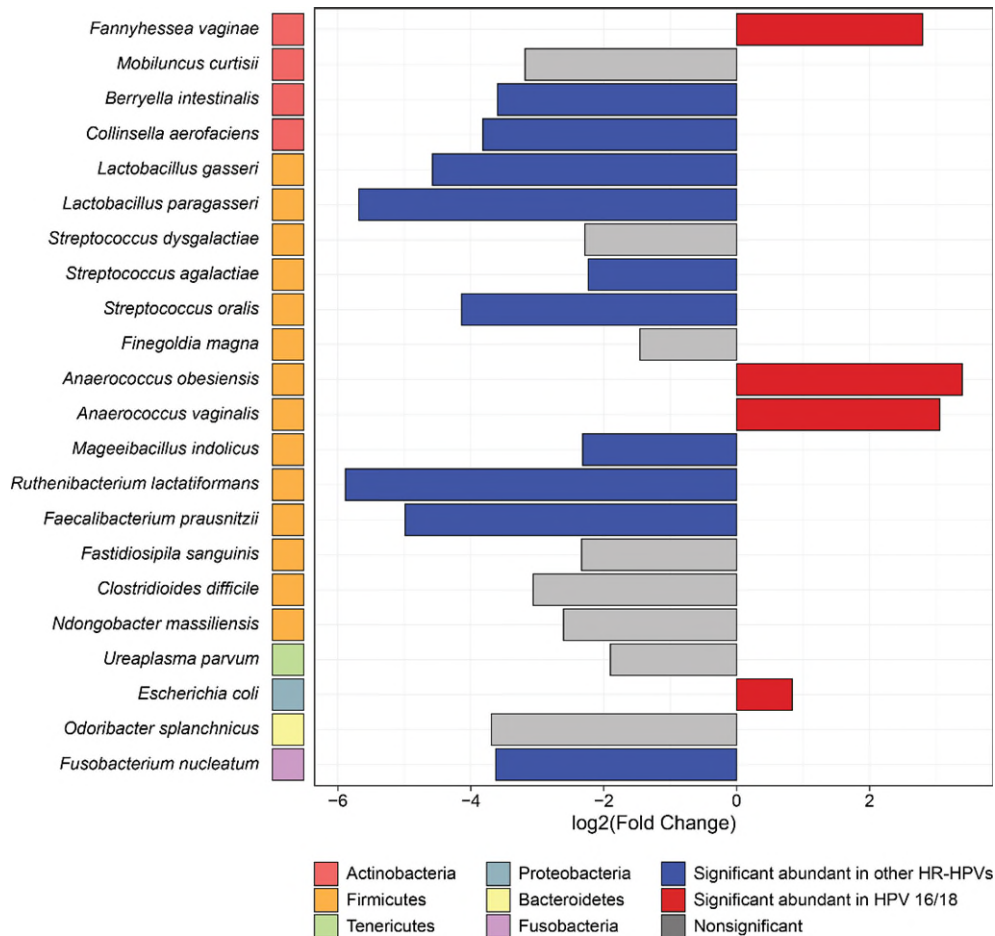
31 **Supplementary Figure 1.** Validation of taxonomic classification using the Human Vaginal Microbiome Genome
 32 Collection (VMGC) in short-read level. Venn diagram showing species (a) and genus-level (b) overlap between
 33 our dataset (classified using Kraken2) and the VMGC reference database. Scatter plots depicting species (c) and
 34 genus-level (d) correlation between Kraken2 and VMGC for vaginal key taxa. Each point represents a sample,
 35 and the blue line represents the linear regression fit. The Spearson correlation coefficient (R) and p-value are
 36 indicated for each taxon.



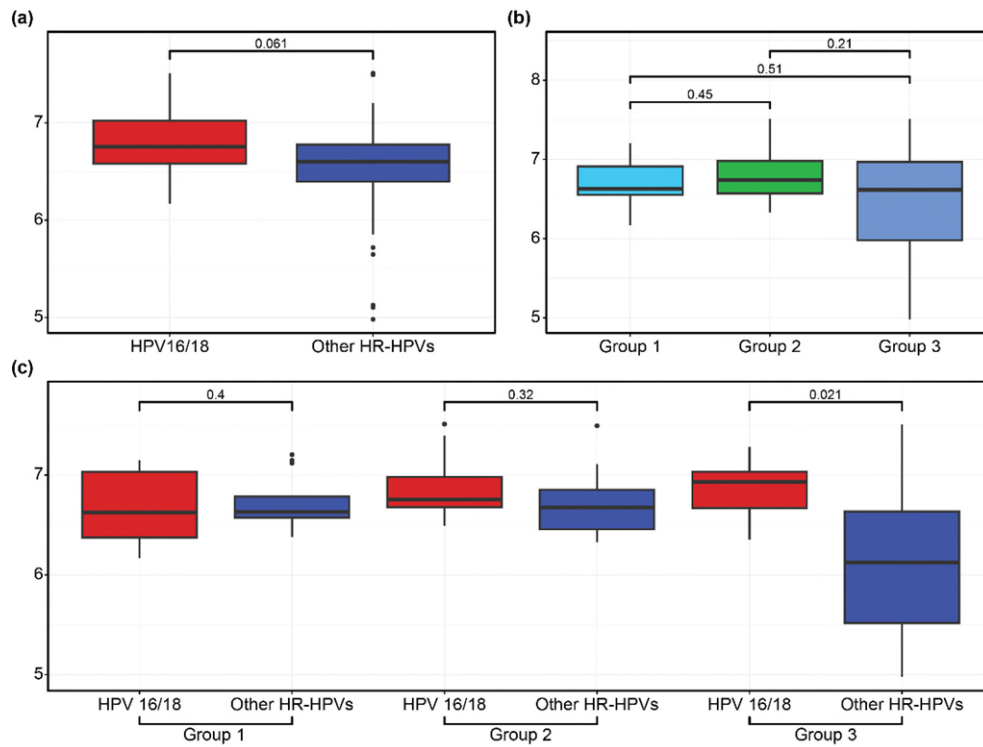
Supplementary Figure 2. Alpha diversity of the vaginal microbial community measured using Shannon index. Comparison of the alpha diversity grouped by: **(a)** HR-HPV types in the entire cohort, **(b)** severity of cervical lesions, **(c)** age range, and **(d)** HR-HPV types across the severity of cervical lesions.



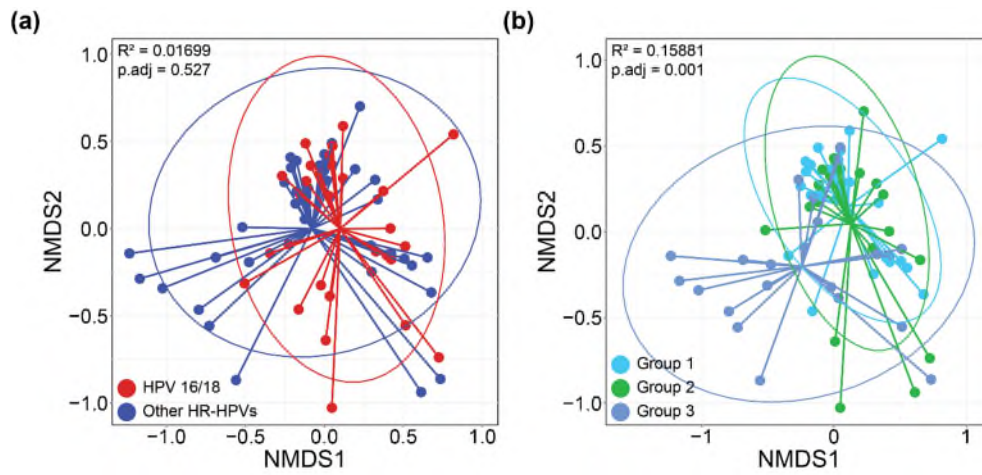
Supplementary Figure 3. Correlation heatmap of bacterial species utilized for vaginal community state types assignment and clinical data. The evaluation of the correlation matrix was performed using Spearman's rank correlation coefficient. Significances are marked by * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$).



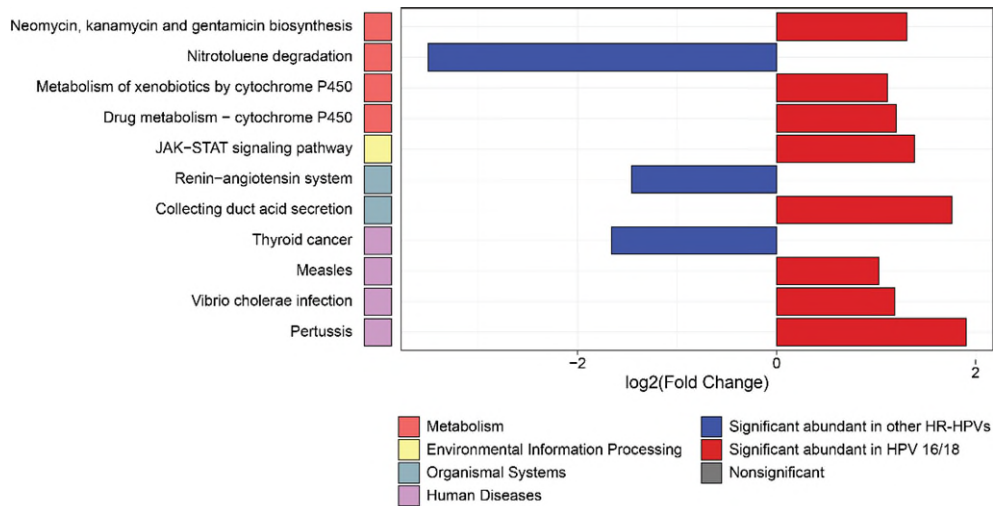
Supplementary Figure 4. Comparison of the vaginal microbiome between HPV 16/18 and other HR-HPVs in the entire cohort. Bar charts display the differentially abundant bacterial species in each disease group (log2 fold change > 1 or < -1, $p < 0.1$, DESeq2), comparing HPV 16/18 and other HR-HPVs. Log2 fold change values are used to denote significant enrichment, with positive values (red bars) indicating enrichment in HPV 16/18 and negative values (blue bars) indicating enrichment in other HR-HPVs (log2 fold change > 1 or < -1, $p < 0.05$, DESeq2).



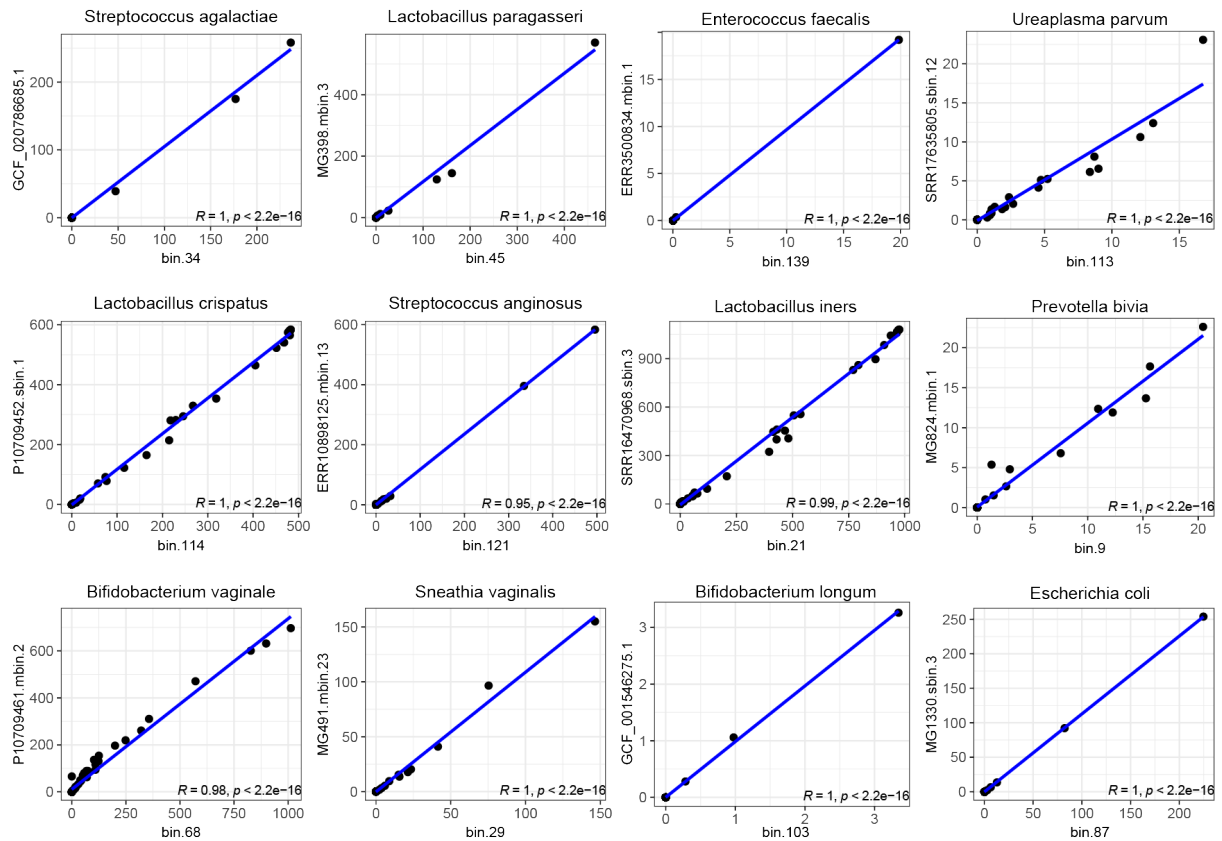
Supplementary Figure 5. Alpha diversity of the functional genes at the KO level measured using Shannon index. Comparison of the alpha diversity grouped by: **(a)** HR-HPV types in the entire cohort, **(b)** severity of cervical lesions, and **(c)** HR-HPV types across the severity of cervical lesions.



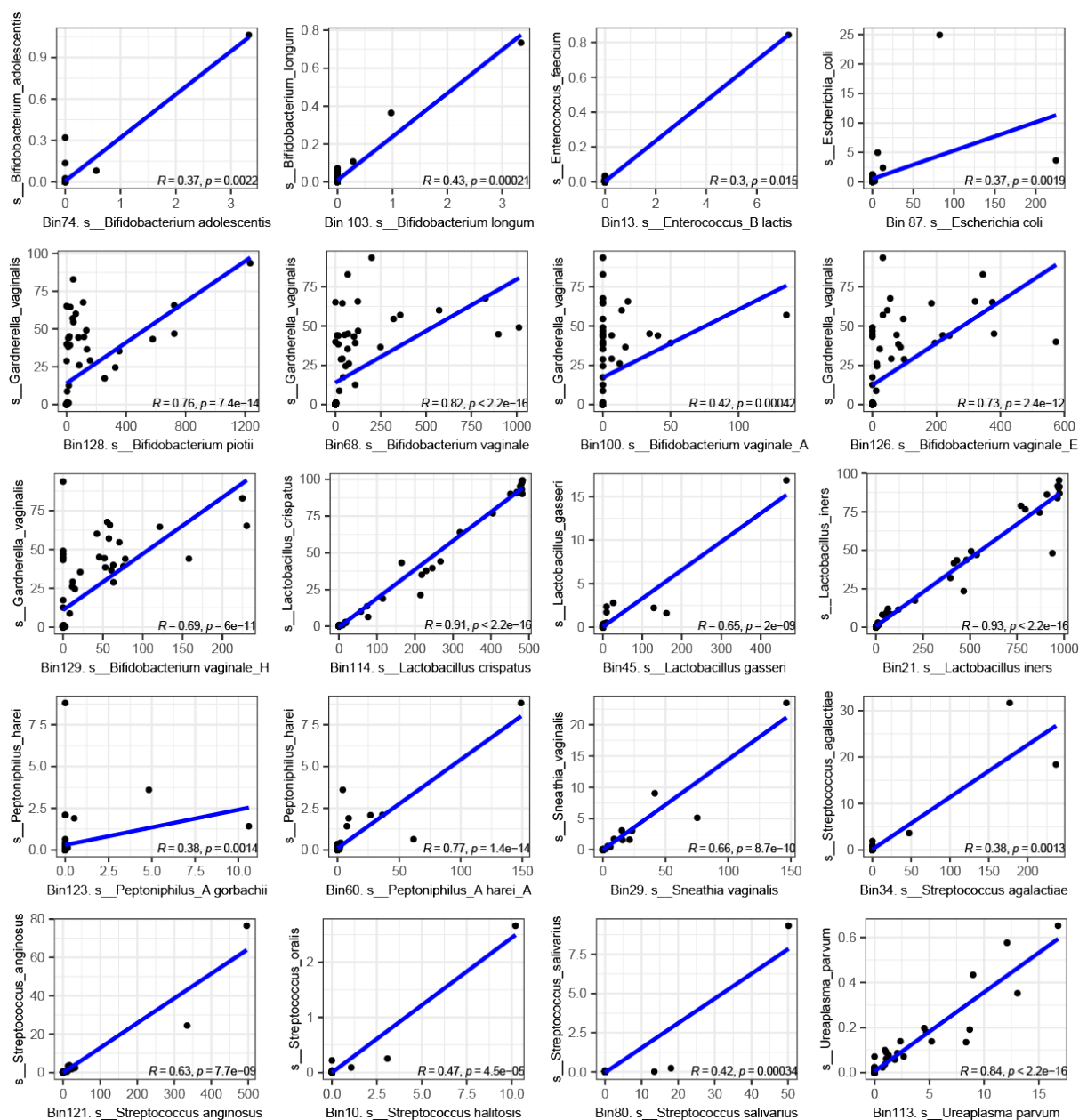
Supplementary Figure 6. Non-metric multidimensional scaling of VM composition in the entire cohort, grouped by **(a)** HR-HPV types and **(b)** severity of cervical lesions.



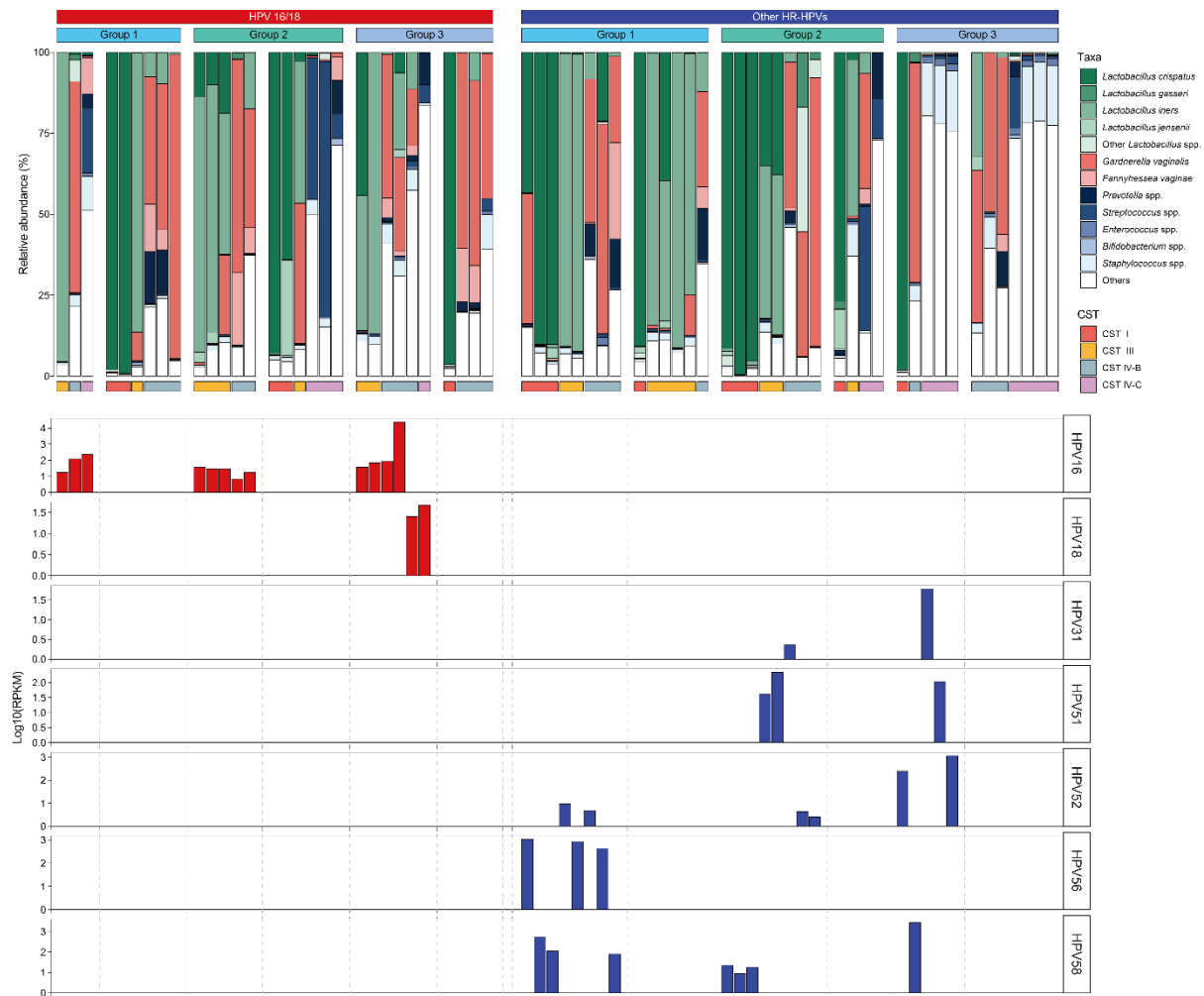
Supplementary Figure 7. Bar charts illustrating differentially abundant KEGG pathways across each disease group (log2 fold change > 1 or < -1, $p < 0.05$, DESeq2) between HPV 16/18 and other HR-HPVs. Log2 fold change values are used to denote significant enrichment, with positive values (red bars) indicating enrichment in HPV 16/18 and negative values (blue bars) indicating enrichment in other HR-HPVs.



Supplementary Figure 8. Validation of taxonomic classification using the Human Vaginal Microbiome Genome Collection (VMGC) in contig level. Scatter plots depicting correlation between MAGs in this study and VMGC for vaginal key taxa. Each point represents a sample, and the blue line represents the linear regression fit. The Spearson correlation coefficient (R) and p-value are indicated for each taxon.



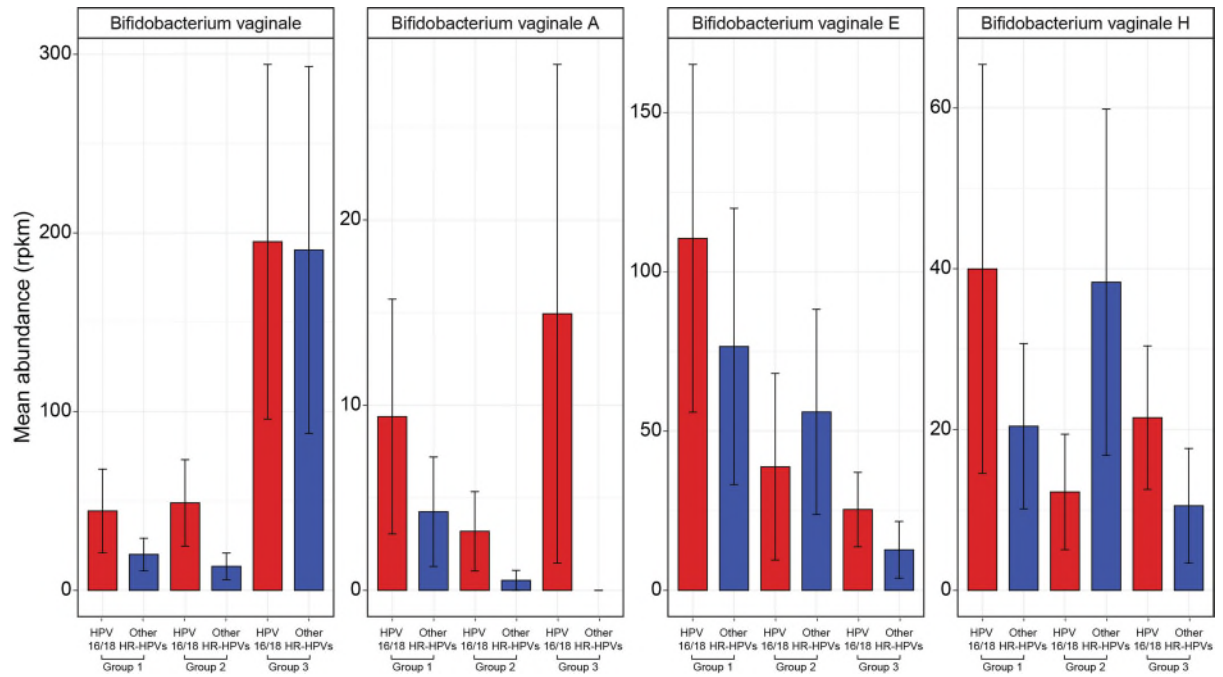
Supplementary Figure 9. Scatter plots comparing the relative abundances of key vaginal microbial taxa identified using contig-based assembly (x-axis) and short-read classification (Kraken2) (y-axis). Each dot represents a sample, and the blue line indicates the linear regression fit. The Spearson correlation coefficient (R) and p-values are provided for each taxon.



Supplementary Figure 10. Vaginal microbiome composition and HPV viral load across study groups. The upper panel shows stacked bar charts of relative microbial abundance in HPV 16/18 and other HR-HPVs groups categorized by the severity of cervical lesions and vaginal community state types (CSTs). The lower panel displays bar plots of HPV RPKM values (log10 scale) for each sample using metagenomic data.



Supplementary Figure 11. Validation of vaginal microbiome composition differences between HPV 16 and other HR-HPV infections using an independent Swedish cohort published by Norenhag et al. Relative abundance of vaginal microbial taxa in the HSIL group, specifically (a) CIN2 and (b) CIN3, comparing HPV 16 infection and other HR-HPV infections.



Supplementary Figure 12. Abundance of four distinct genotypes of *Bifidobacterium vaginalis* bins.

Variables	HPV 16/18 (n = 30)	Other HR-HPVs (n = 38)	<i>p</i> value
Age (years)	40.7 ± 9.4	45.5 ± 10.7	0.0589
BMI (kg m ⁻²)	24.4 ± 3.7	24.9 ± 3.5	0.5213
Married women (n, %)	22 (73.3)	35 (92.1)	0.0383
Parity (n)	1.3 ± 1.2	1.7 ± 1.0	0.1524
Smoking (n, %)	7 (23.3)	3 (7.9)	0.0764
Disease status			
Group 1	9 (30.0)	14 (36.8)	
Group 2	11 (36.7)	12 (31.6)	
Group 3	10 (33.3)	12 (31.6)	
HPV type (n, %)			< 0.0001
16	25 (83.3)		
18	5 (16.7)		
31		3 (7.9)	
33		2 (5.3)	
36		1 (2.6)	
39		1 (2.6)	
45		2 (5.3)	
51		6 (15.8)	
52		10 (26.3)	

	53		1 (2.6)	
	56		3 (7.9)	
	58		10 (26.3)	
	62		1 (2.6)	
	Multiple HPV infection (n, %)	0 (0)	2 (5.3)	0.4995
99				
100				

101 **Supplementary Table 2.** Distribution of vaginal community state types (CSTs) in the entire cohort

CST	HPV 16/18 (n = 30)	Other HR-HPVs (n = 38)	<i>p</i> value
I + III (n, %)	13 (43.3%)	18 (47.4%)	0.7446
I	5 (16.7%)	9 (23.7%)	0.4781
I-A	4 (13.3%)	7 (18.4%)	0.5725
I-B	1 (3.3%)	2 (5.3%)	0.6984
III	8(26.7%)	9 (23.7%)	0.783
III-A	5 (16.7%)	5 (13.2%)	0.694
III-B	3(10.0%)	4 (10.5%)	0.9444
IV	17 (56.7%)	20 (52.6%)	0.7446
IV-B	12 (40.0%)	11 (28.9%)	0.3512
IV-C0	2 (6.7%)		0.1608
IV-C1	3 (10.0%)	3 (7.9%)	0.7685
IV-C4		6 (15.8%)	< 0.05

102

103

	Entire cohort		Group 1		Group 2		Group 3	
	r ²	<i>p</i> value	r ²	<i>p</i> value	r ²	<i>p</i> value	r ²	<i>p</i> value
<i>Lactobacillus crispatus</i>	0.373	0.001	0.309	0.033	0.274	0.029	0.606	0.004
<i>Lactobacillus iners</i>	0.088	0.042	0.295	0.028	0.132	0.244	0.035	0.673
<i>Gardnerella vaginalis</i>	0.213	0.002	0.549	0.003	0.484	0.005	0.072	0.500
<i>Fannyhessea vaginae</i>	0.230	0.002	0.517	0.010	0.219	0.118	0.369	0.028
<i>Streptococcus</i> spp.	0.080	0.076	0.033	0.661	0.206	0.092	0.551	0.004
<i>Staphylococcus</i> spp.	0.337	0.001	0.153	0.193	0.348	0.026	0.604	0.001

Supplementary Table 4. The genomic assembly characteristics of MAGs.

MAG	Genome size (Mb)	Completeness (%)	Contamination (%)	MIMIG	Taxa
bin.1	2.17	96.72	1.67	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Pr evotella;s_Prevotella sp000758925
bin.2	1.74	91.49	1.05	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Peptostreptococcales;f_Filifactoraceae; g_Criibacterium;s_Criibacterium bergeronii
bin.3	1.73	84.99	0.95	Medium quality	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Actinomyc etaceae;g_Mobiluncus;s_
bin.4	1.52	86.79	2.13	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_He lcococcus;s_Helcococcus sueciensis
bin.5	1.97	77.55	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_ Ruminococcus_D;s_Ruminococcus_D bicirculans
bin.6	2.13	97.99	1.44	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Rikenellaceae;g_Ali stipes;s_Alistipes putredinis
bin.7	1.46	82.56	1.17	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Oscillospirales;f_Oscillospiraceae;g_E R4;s_ER4 sp900317525
bin.8	1.16	58.64	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_An aerococcus;s_Anaerococcus obesiensis
bin.9	1.97	99.31	0.00	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Pr evotella;s_Prevotella bivia
bin.10	1.73	87.26	1.51	Medium quality	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g_Strepto coccus;s_Streptococcus halitosis
bin.11	2.24	95.85	0.95	Near complete	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Actinomyc etaceae;g_Mobiluncus;s_Mobiluncus mulieris
bin.12	1.26	94.4	1.27	Near complete	d_Bacteria;p_Bacillota_C;c_Negativicutes;o_Veillonellales;f_Dialisteraceae;g_D ialister;s_Dialister sp001553355
bin.13	2.52	98.49	0.38	Near complete	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Enterococcaceae;g_Enteroc occus_B;s_Enterococcus_B lactis

bin.14	1.23	93.67	0.00	Near complete	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Actinomycetaceae;g_Arcanobacterium_A;s_Arcanobacterium_A urinimassiliense
bin.15	1.31	62	0.89	Medium quality	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Prevotella;s_
bin.16	1.24	99.04	0.00	Near complete	d_Bacteria;p_Bacillota;c_Bacilli;o_Erysipelotrichales;f_Erysipelotrichaceae;g_Bulleidia;s_
bin.17	1.36	73.61	3.69	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Oscillospirales;f_Oscillospiraceae;g_Vescimonas;s_Vescimonas sp900545585
bin.18	1.33	76.7	1.03	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Peptostreptococcales;f_Anaerovoracaceae;g_Mogibacterium;s_
bin.19	1.54	97.58	0.12	Near complete	d_Bacteria;p_Actinomycetota;c_Coriobacteriia;o_Coriobacteriales;f_Eggerthellaaceae;g_Berryella;s_Berryella sp900604965
bin.20	2.76	92.3	1.34	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Coprococcus_A;s_Coprococcus_A catus_A
bin.21	1.03	93.83	0.77	Near complete	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Lactobacillaceae;g_Lactobacillus;s_Lactobacillus iners
bin.22	1.71	94.3	1.01	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Porphyromonadaceae;g_Porphyromonas;s_Porphyromonas sp001552775
bin.23	1.62	53.57	0.89	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_Faecalibacterium;s_Faecalibacterium prausnitzii_D
bin.24	1.52	100	0.10	Near complete	d_Bacteria;p_Bacillota_C;c_Negativicutes;o_Veillonellales;f_Megasphaeraceae;g_Megasphaera;s_Megasphaera lornae
bin.25	6.78	96.26	4.31	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Eisenbergiella;s_Eisenbergiella tayi
bin.26	1.30	67.7	0.17	Medium quality	d_Bacteria;p_Bacillota_B;c_Peptococcia;o_Peptococcales;f_Peptococcaceae;g_Peptococcus;s_Peptococcus niger
bin.27	1.52	91.25	1.75	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Ezakiella;s_Ezakiella coagulans
bin.28	1.67	86.28	0.10	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Anaerococcus;s_Anaerococcus degeneri

bin.29	1.13	97.75	2.81	Near complete	d_Bacteria;p_Fusobacteriota;c_Fusobacteriia;o_Fusobacteriales;f_Leptotrichiaceae;g_Sneathia;s_Sneathia vaginalis
bin.30	1.46	50.6	1.88	Medium quality	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Actinomycetaceae;g_Varibaculum;s_Varibaculum massiliense
bin.31	1.54	92.48	1.75	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Peptoniphilus_A;s_Peptoniphilus_A grossensis
bin.33	1.72	99.58	0.00	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Peptostreptococcales;f_Peptostreptococaceae;g_Peptostreptococcus;s_Peptostreptococcus anaerobius
bin.34	2.38	100	1.86	Near complete	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g_Streptococcus;s_Streptococcus agalactiae
bin.35	2.18	73.52	1.79	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_Faecalibacterium;s_Faecalibacterium prausnitzii
bin.36	1.36	52.75	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Anaerococcus;s_Anaerococcus senegalensis
bin.37	2.03	78.13	3.99	Medium quality	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Porphyrimonadaceae;g_Porphyrimonas;s_Porphyrimonas sp900538385
bin.38	2.35	98.64	0.39	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Prevotella;s_Prevotella corporis
bin.39	1.19	93.93	0.09	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Parvimonas;s_Parvimonas sp001552895
bin.40	2.43	74.24	5.35	Medium quality	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Prevotella;s_Prevotella stercora
bin.41	0.64	100	0.38	Near complete	d_Bacteria;p_Bacillota;c_Bacilli;o_Mycoplasmatales;f_Metamycoplasmataceae;g_Metamycoplasma;s_Metamycoplasma hominis
bin.42	1.52	96.24	0.00	Near complete	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Aerococcaceae;g_Aerococcus;s_Aerococcus christensenii
bin.43	3.12	86.25	0.63	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Blautia_A;s_Blautia_A wexlerae
bin.44	1.32	96.5	0.77	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Peptoniphilus_A;s_Peptoniphilus_A lacrimalis

bin.45	1.75	97.8	0.09	Near complete	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Lactobacillaceae;g_Lactobacillus;s_Lactobacillus paragasseri
bin.46	1.14	97.25	0.22	Near complete	d_Bacteria;p_Bacillota_C;c_Negativicutes;o_Veillonellales;f_Dialisteraceae;g_Dialister;s_Dialister micraerophilus
bin.47	1.05	62.68	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Anaerococcus;s_Anaerococcus lactolyticus
bin.48	1.51	96.55	0.00	Near complete	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Actinomycetaceae;g_Arcanobacterium_A;s_Arcanobacterium_A sp000758825
bin.49	1.39	87.28	0.58	Medium quality	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Lactobacillaceae;g_Limosilactobacillus;s_Limosilactobacillus coleohominis_A
bin.50	1.89	97.31	0.00	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Oscillospirales;f_Acutalibacteraceae;g_CAG-217;s_CAG-217 sp000436335
bin.51	2.10	57.89	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Oliverpabstia;s_Oliverpabstia intestinalis
bin.52	1.00	66.13	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Peptoniphilus_C;s_Peptoniphilus_C pacaensis
bin.54	2.13	64.03	1.75	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_KLE1615;s_KLE1615 sp900066985
bin.55	0.58	64.65	0.00	Medium quality	d_Bacteria;p_Bacillota_C;c_Negativicutes;o_Veillonellales;f_Dialisteraceae;g_Dialister;s_Dialister sp900541485
bin.56	2.30	96.22	0.50	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Muribaculaceae;g_Sodaliophilus;s_
bin.57	1.39	89.59	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Saccharofermentanales;f_Fastidiosipila ceae;g_KA00274;s_KA00274 sp001552885
bin.58	1.84	98.36	1.39	Near complete	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Lactobacillaceae;g_Limosilactobacillus;s_Limosilactobacillus portuensis
bin.59	0.99	83.65	1.61	Medium quality	d_Bacteria;p_Actinomycetota;c_Coriobacteriia;o_Coriobacteriales;f_Atopobiaceae;g_Fannyhessea;s_
bin.60	1.04	71.15	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Peptoniphilus_A;s_Peptoniphilus_A harei_A

bin.61	1.62	98.78	0.30	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Anaerococcus;s_Anaerococcus_tetradus
bin.62	2.57	97.15	1.09	Near complete	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Mycobacteriales;f_Mycobacteriaceae;g_Corynebacterium;s_Corynebacterium_pyruviciproducens
bin.63	1.21	83.68	0.67	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Peptoniphilus_B;s_Peptoniphilus_B_duerdenii
bin.65	1.67	98.87	0.00	Near complete	d_Bacteria;p_Fusobacteriota;c_Fusobacteriia;o_Fusobacteriales;f_Fusobacteriaceae;g_Fusobacterium_C;s_Fusobacterium_C_gonidiaformans
bin.66	2.12	99.47	2.88	Near complete	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Lactobacillaceae;g_Ligilactobacillus;s_Ligilactobacillus_salivarius
bin.67	1.34	80.08	0.75	Medium quality	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Actinomycetaceae;g_Pauljensenia;s_Pauljensenia_turicensis
bin.68	0.91	65.27	3.69	Medium quality	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Bifidobacteriaceae;g_Bifidobacterium;s_Bifidobacterium_vaginale
bin.69	2.19	98.31	0.34	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Prevotella;s_Prevotella_disiens
bin.70	1.62	90.48	0.00	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Porphyromonadaceae;g_Porphyromonas;s_Porphyromonas
bin.71	1.89	93.26	1.75	Near complete	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Actinomycetaceae;g_Varibaculum;s_Varibaculum_cambriense_A
bin.72	1.43	58.53	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Dorea;s_Dorea_formicigenerans
bin.73	1.22	91.48	0.00	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Peptostreptococcales;f_Anaerovoracaceae;g_S5-A14a;s_S5-A14a_sp000758905
bin.74	1.87	90.14	1.28	Near complete	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Bifidobacteriaceae;g_Bifidobacterium;s_Bifidobacterium_adolescentis
bin.75	1.53	65.22	3.68	Medium quality	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Prevotella;s_Prevotella_sp000479005
bin.77	1.08	83.36	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Finegoldia;s_Finegoldia_magna_H

bin.78	3.91	96.14	0.19	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Phocaeicola;s_Phocaeicola vulgatus
bin.79	1.47	95.73	0.00	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Levyella;s_Levyella massiliensis
bin.80	1.23	70.19	0.78	Medium quality	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g_Streptococcus;s_Streptococcus salivarius
bin.81	1.82	96.62	0.63	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Stomatobaculum;s_Stomatobaculum sp002892395
bin.84	1.91	98.59	0.07	Near complete	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Lactobacillaceae;g_Limosilactobacillus;s_Limosilactobacillus reuteri_E
bin.85	1.18	72.63	3.19	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Peptoniphilus_B;s_Peptoniphilus_B sp000478985
bin.86	1.60	94.76	0.24	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_UBA932;g_Cryptobacteroides;s_
bin.87	4.45	98.58	0.23	Near complete	d_Bacteria;p_Pseudomonadota;c_Gammaproteobacteria;o_Enterobacterales;f_Enterobacteriaceae;g_Escherichia;s_Escherichia coli
bin.88	1.77	94.92	0.00	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Prevotella;s_
bin.89	2.04	71.67	4.82	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_Gemmiger;s_Gemmiger sp900539695
bin.90	1.89	86.53	0.85	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_Faecalibacterium;s_Faecalibacterium longum
bin.91	1.48	96.85	0.04	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Peptoniphilus_C;s_Peptoniphilus_C coxii
bin.92	1.27	68.31	1.72	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Anaerococcus;s_Anaerococcus vaginalis_B
bin.93	1.98	98.65	1.05	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Oscillospirales;f_Acutalibacteraceae;g_Ruminococcus_E;s_Ruminococcus_E bromii_B
bin.94	1.47	82.58	0.52	Medium quality	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Actinomycetaceae;g_Varibaculum;s_Varibaculum timonense_A

bin.95	3.29	78.22	2.87	Medium quality	d_Bacteria;p_Bacillota;c_Bacilli;o_Erysipelotrichales;f_Copro bacillaceae;g_Coprobacillus;s_Copro bacillus cateniformis
bin.96	2.14	60.52	4.39	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Agathobacter;s_Agathobacter faecis
bin.97	1.58	93.73	0.71	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Saccharofermentanales;f_Fastidiosipila ceae;g_Mageeibacillus;s_Mageeibacillus indolicus
bin.98	1.62	82.27	1.01	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Oscillospirales;f_Oscillospiraceae;g_Vescimonas;s_Vescimonas coprocola
bin.99	0.75	57.26	1.14	Medium quality	d_Bacteria;p_Patescibacteria;c_Saccharimonadia;o_Saccharimonadales;f_Nanop eriomorbaseae;g_Nanoperiomorbus;s_Nanoperiomorbus sp004136275
bin.100	1.16	82.32	1.91	Medium quality	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Bifidobact eriaceae;g_Bifidobacterium;s_Bifidobacterium vaginale_A
bin.101	2.49	98.34	0.00	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Mediterraneibacter;s_Mediterraneibacter torques
bin.102	1.33	77.24	1.93	Medium quality	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Porphyrimonadacea e;g_Porphyrimonas;s_
bin.103	1.63	85.17	2.11	Medium quality	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Bifidobact eriaceae;g_Bifidobacterium;s_Bifidobacterium longum
bin.104	2.24	97.46	0.00	Near complete	d_Bacteria;p_Bacillota_C;c_Negativicutes;o_Selenomonadales;f_Selenomonada ceae;g_Megamonas;s_Megamonas funiformis
bin.105	2.03	84.27	2.27	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_Faecalibacterium;s_Faecalibacterium duncaniae
bin.107	1.29	88.73	0.35	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Pe ptoniphilus_C;s_Peptoniphilus_C sp902363535
bin.108	1.58	95.97	0.00	Near complete	d_Bacteria;p_Bacillota_C;c_Negativicutes;o_Veillonellales;f_Dialisteraceae;g_A llisonella;s_Allisonella pneumosinta
bin.109	2.48	55.59	0.00	Medium quality	d_Bacteria;p_Bacillota;c_Bacilli;o_Erysipelotrichales;f_Erysipelotrichaceae;g_C lostridium_AQ;s_Clostridium_AQ innocuum
bin.110	1.55	76.2	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_UBA629;s_UBA629 sp005465875

bin.111	1.28	86.4	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Saccharofermentanales;f_Fastidiosipila ceae;g_KA00274;s_KA00274 sp902373515
bin.112	0.99	72.16	0.42	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Fe nollaria;s_Fenollaria massiliensis
bin.113	0.72	99.04	0.00	High quality	d_Bacteria;p_Bacillota;c_Bacilli;o_Mycoplasmatales;f_Mycoplasmoidaceae;g_ Ureaplasma;s_Ureaplasma parvum
bin.114	2.07	97.29	0.15	Near complete	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Lactobacillaceae;g_Lactob acillus;s_Lactobacillus crispatus
bin.115	1.31	63.23	0.67	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_ Ruminococcus_C;s_Ruminococcus_C sp000433635
bin.116	1.12	84.59	1.29	Medium quality	d_Bacteria;p_Campylobacterota;c_Campylobacteria;o_Campylobacteriales;f_Ca mpylobacteraceae;g_Campylobacter_B;s_Campylobacter_B ureolyticus
bin.117	1.37	91.06	0.59	Near complete	d_Bacteria;p_Actinomycetota;c_Coriobacteriia;o_Coriobacteriales;f_Eggerthella ceae;g_Berryella;s_Berryella sp001552935
bin.118	1.22	93.81	4.37	Near complete	d_Bacteria;p_Actinomycetota;c_Coriobacteriia;o_Coriobacteriales;f_Atopobiace ae;g_Fannyhessea;s_Fannyhessea massiliensis
bin.119	4.04	96.29	0.81	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_B acteroides;s_Bacteroides caccae
bin.120	1.89	98.98	0.00	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Pr evotella;s_Prevotella amnii
bin.121	1.86	96.79	2.54	High quality	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g_Strepto coccus;s_Streptococcus anginosus
bin.122	1.18	94.61	1.12	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Par vimonas;s_Parvimonas parva
bin.123	1.12	75.3	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Pe ptoniphilus_A;s_Peptoniphilus_A gorbachii
bin.124	2.63	96.97	2.22	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_F usicatenibacter;s_Fusicatenibacter saccharivorans
bin.125	1.67	93.59	0.71	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Peptostreptococcales;f_Anaerovoracac eae;g_S5-A14a;s_S5-A14a sp900553025

bin.126	0.95	57.68	5.17	Medium quality	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Bifidobacteriaceae;g_Bifidobacterium;s_Bifidobacterium_vaginale_E
bin.127	1.62	98.51	0.00	Near complete	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Peptostreptococcales;f_Peptostreptococcaceae;g_Peptostreptococcus;s_Peptostreptococcus_sp000758885
bin.128	0.63	50.52	3.60	Medium quality	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Bifidobacteriaceae;g_Bifidobacterium;s_Bifidobacterium_piotii
bin.129	1.48	90.98	4.78	Near complete	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Bifidobacteriaceae;g_Bifidobacterium;s_Bifidobacterium_vaginale_H
bin.130	1.10	84.88	0.95	Medium quality	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Actinomycetaceae;g_ZJ293;s_
bin.131	1.02	61.94	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Tissierellales;f_Peptoniphilaceae;g_Fenollaria;s_
bin.132	4.15	97.27	1.32	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Bacteroides;s_Bacteroides_uniformis
bin.133	1.37	100	0.00	Near complete	d_Bacteria;p_Actinomycetota;c_Coriobacteriia;o_Coriobacteriales;f_Atopobiaceae;g_Atopobium;s_Atopobium_deltae
bin.134	1.87	95	0.00	Near complete	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Mycobacteriales;f_Mycobacteriaceae;g_Lawsonella;s_Lawsonella_clevelandensis
bin.135	1.64	97.28	0.32	Near complete	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Lactobacillaceae;g_Lactobacillus;s_Lactobacillus_mulieris
bin.136	3.53	97.63	0.00	Near complete	d_Bacteria;p_Desulfobacterota;c_Desulfovibrionia;o_Desulfovibrionales;f_Desulfovibrionaceae;g_Bilophila;s_Bilophila_wadsworthia
bin.137	1.48	61.6	0.00	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_Gemmiger;s_Gemmiger_formicilis
bin.138	2.15	77.27	0.53	Medium quality	d_Bacteria;p_Bacillota_A;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Blautia_A;s_Blautia_A_massiliensis
bin.139	2.90	98.11	1.03	Near complete	d_Bacteria;p_Bacillota;c_Bacilli;o_Lactobacillales;f_Enterococcaceae;g_Enterococcus;s_Enterococcus_faecalis
bin.140	1.24	68.29	1.10	Medium quality	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Porphyromonadaceae;g_Porphyromonas_A;s_Porphyromonas_A_sp001808555

bin.141	3.62	96.11	2.30	Near complete	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_P hocaeicola;s_Phocaeicola coprocola
bin.142	1.24	98.31	0.16	Near complete	d_Bacteria;p_Fusobacteriota;c_Fusobacteriia;o_Fusobacteriales;f_Leptotrichiace ae;g_Sneathia;s_Sneathia sanguinegens
bin.143	2.32	97.78	8.38	Medium quality	d_Bacteria;p_Actinomycetota;c_Actinomycetia;o_Actinomycetales;f_Actinomyc etaceae;g_Mobiluncus;s_
bin.144	3.80	67.86	5.11	Medium quality	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_B acteroides;s_Bacteroides thetaiotaomicron

107

108

109 **Supplementary Table 5.** Comparison of HPV detection between metagenomic sequencing and qPCR across study samples.

		Metagenome HPV-positive / qPCR HPV-positive (Number of samples)	Metagenome HR-HPV-positive / qPCR HR-HPV-positive (Number of samples)	Metagenome HPV 16/18-positive / qPCR HPV 16/18-positive (Number of samples)	Metagenome other HR-HPVs-positive / qPCR other HR-HPVs-positive (Number of samples)
Total	HPV 16/18	15 / 30	14 / 30	14 / 30	0 / 0
	Other HR-HPVs	22 / 38	21 / 38	0 / 0	21 / 38
Group 1	HPV 16/18	4 / 9	3 / 9	3 / 9	0 / 0
	Other HR-HPVs	9 / 14	8 / 14	0 / 0	8 / 14
Group 2	HPV 16/18	5 / 11	5 / 11	5 / 11	0 / 0
	Other HR-HPV	8 / 12	8 / 12	0 / 0	8 / 12
Group 3	HPV 16/18	6 / 10	6 / 10	6 / 10	0 / 0
	Other HR-HPVs	5 / 12	5 / 12	0 / 0	5 / 12