

<https://doi.org/10.1038/s41522-025-00690-1>

Navigating complexities of polymorphic microbiomes in endometrial cancer



Nicole R. Jimenez¹, Chloe R. Herman^{2,3}, Paweł Łaniewski⁴, Emily Cope², Keehoon Lee⁵,
Nichole D. Mahnert¹, Dana M. Chase⁶, J. Gregory Caporaso^{2,4}✉ & Melissa M. Herbst-Kralovetz^{1,4}✉

The microbiome is key to understanding endometrial cancer (EC) etiology and prevention strategies, implicated in the regulation of estrogen in estrogen-driven cancers. Utilizing robust methodologies in the QIIME 2 platform, we examined 16S rRNA vaginal and rectal microbiome data from an EC cohort: 192 women with benign gynecologic conditions, endometrial hyperplasia, or endometrial cancer. Distinct microbial compositions and community networks specific to EC were identified and related to histological grade with adjustments for EC risk factors. Vaginal health-associated *Lactobacillus* and *Limosilactobacillus*, and rectal *Prevotella* and *Peptoniphilus*, were depleted in EC, while detrimental vaginal *Anaerococcus*, *Porphyromonas*, *Prevotella*, *Peptoniphilus*, and rectal *Buttiauxella* were enriched. Significant bacterial features were shared between rectal and vaginal sites in EC, such as *Prevotella timonensis* and *Peptoniphilus A*. Vaginal *Lactobacillus* abundance contributed to less feature sharing from the rectum. Putative microbial metabolic analysis identified dysregulation of amino acid, complex carbohydrate, and hormone metabolism amongst patients with EC.

Endometrial cancer (EC) is the fourth most common female cancer in high-income countries^{1,2}, and it is predicted that by 2030, there will be 122,000 new cases per year in the US³. However, low and middle-income countries have a higher risk of mortality due to delay in diagnosis, access to care, and a higher proportion of aggressive non-endometrioid tumors⁴. Recent findings indicate known disparities in mortality rates by race/ethnicity, with an alarming increase in mortality over time for Black patients specifically⁵. Further an area for improvement by innovative methods is to decrease delay in care which is seen in Hispanic and Black populations⁵.

EC is a cancer of the inner epithelial lining of the uterus, the endometrium⁶. Historically, histopathology from endometrial biopsies was utilized for EC subtyping into types I and II. Type I EC develops through precursor lesions known as endometrial intraepithelial neoplasia; this type is composed of grade 1 or 2 endometrioid carcinoma and has been associated with estrogen dysregulation or, more specifically, hyperestrogenism⁷. On the other hand, type II EC is less common and is classified by higher-grade endometrioid carcinomas or other non-endometrioid subtypes. Type II EC is more aggressive, has a poorer prognosis, and arises due to genotoxic stress⁷. More recently, molecular categorization has been investigated based on gene mutation burden or copy number alterations from patients with EC. These molecular investigations have yielded additional details on the

prognosis of disease and guides personalized medical care in terms of treatment⁸.

Further, the well-established *Hallmarks of Cancer*, a heuristic model created to better define the pathophysiology of cancer, has been investigated for novel or more efficient treatments of cancers⁹. In 2022, four emerging *Hallmarks* and *Enabling Characteristics* were proposed for this model, including polymorphic microbiomes¹⁰. Bacteria inhabiting the female reproductive tract may contribute to gynecologic cancer development by modulating cancer hallmarks such as inflammation, genomic instability, angiogenesis, epithelial barrier disruption, and metabolic dysregulation¹¹.

There is support that the human microbiome may contribute to up to 15% of cancers¹². This is unsurprising, as bacteria such as *Helicobacter pylori*¹³, *Fusobacterium*¹⁴, and *Porphyromonas*¹⁵ have been directly associated with carcinogenesis in the GI tract. Cervicovaginal bacteria such as *Sneathia*¹⁶, *Anaerococcus*¹⁷, *Chlamydia*¹⁸, *Atopobium/Fannyhessea*¹⁹, and others²⁰ are currently under investigation for their indirect or direct role in gynecologic cancers. In the vagina, the genus *Lactobacillus* plays a crucial role in the maintenance of homeostasis and may be protective against gynecologic cancers²¹. Steroid hormones are implicated in some cancers, such as EC, and also impact the vaginal microbiome^{22,23}. The interplay of microbiomes from multiple body sites is referred to as microbiome axes.

¹Department of Obstetrics and Gynecology, College of Medicine-Phoenix, University of Arizona, Phoenix, AZ, USA. ²Center for Applied Microbiome Science, Pathogen and Microbiome Institute, Northern Arizona University, Flagstaff, AZ, USA. ³School of Informatics, Computing and Cyber Systems, Northern Arizona University, Flagstaff, AZ, USA. ⁴Department of Basic Medical Sciences, College of Medicine-Phoenix, University of Arizona, Phoenix, AZ, USA. ⁵Translational Genomics Research Institute, part of City of Hope, Flagstaff, AZ, USA. ⁶David Geffen School of Medicine at UCLA, Los Angeles, CA, USA.

✉ e-mail: greg.caporaso@nau.edu; mherbst1@arizona.edu

There is interest in investigating microbiome axes' contributions to gynecologic conditions, such as through the gut-vagina axis²⁴. This axis is where the gut/rectal microbiome^{25,26} plays a role in estrogen metabolism, which has been linked to conditions like EC in vaginal and endometrial sites^{17,27}. We hypothesize that investigating microbiomes at multiple mucosal sites will enable a better understanding of the potential role of microorganisms in EC development and etiology.

Previous studies have identified microbial associations with endometrial cancer^{19,28–30}; however, interpretation of the findings can be challenging due to differences in study design and failure to account for confounding factors that affect microbial composition. Known risk factors of EC include diagnosis of endometrial hyperplasia, older age, menopause, higher BMI, estrogen exposure, tamoxifen usage, low parity, metabolic syndrome, and Lynch syndrome (a genetic predisposition for certain cancers)^{7,31}. Since the microbiome can be affected by factors such as age, hormone status, BMI, and co-morbidities, multiple clinical studies are needed to validate the clinical linkage to microbial composition of EC and not signatures related to underlying risk factors.

Herein, we tested our hypothesis that the microbiomes of both the rectum and vagina will reveal distinct microbial signatures between patients newly diagnosed with EC and benign gynecologic conditions. To investigate this, we collected vaginal and rectal swabs from 192 patients undergoing hysterectomy and diagnosed with EC, hyperplasia, or benign conditions and performed robust and novel microbiome evaluations, which included adjustments for known risk factors of EC. This is an extensive EC microbiome study investigating the interplay of proximal and local microbiomes, specifically the gut-vagina axis, in patients newly diagnosed with EC. We present global microbiome states associated with endometrial cancer, significant microbial signatures after adjustment for known risk factors, differences in microbial community structures, and metabolic potential of microbiota in EC compared to benign conditions. This study identifies putative oncobacteria, bacteria with pro-oncogenic effects, for further investigation that may have pathophysiologic roles as either drivers of carcinogenesis or passengers that thrive in an endometrial cancer environment. Overall, unique microbial features identified in this study will allow for future investigation into risk reduction and prevention of EC.

Results

Patient characteristics

The patient population for this study consisted of 192 women undergoing hysterectomy for newly diagnosed cancer or benign conditions. Following surgery, histopathology confirmed the patients' diagnoses. Participants were grouped as follows: grade 1/2 endometrial endometrioid carcinoma (EEC) ($n = 53$); other endometrial cancer subtypes (other ECs) ($n = 13$); endometrial hyperplasia ($n = 18$); and benign conditions ($n = 108$) (Table 1). Age significantly differed between benign and all other groups (p value 0.007, <0.0001, <0.0001, respectively) (Table 1). EC is most often diagnosed in postmenopausal women, women with a higher BMI, and women undergoing procedures such as dilation and curettage (D&C), which may disrupt the uterine lining. In alignment with these risk factors, our study identified that postmenopausal status was significantly different between benign conditions (17.59%) and hyperplasia (66.67%), grade 1/2 EEC (76.47%), other ECs (92.31%) (p value < 0.0001 for all malignant groups) (Table 1). BMI was also significantly different between women who were diagnosed with benign conditions, hyperplasia, grade 1/2 EEC, and other ECs (p value < 0.0001) (Table 1). Additionally, having undergone a D&C differed between benign conditions (17.59%), hyperplasia (38.89%), grade 1/2 EEC (47.17%), and other ECs (53.85%) (Table 1). Further, hormonal contraceptives were also different between groups (p value = 0.01), where 32.1% of patients using hormonal contraceptives were diagnosed with benign conditions, which may be due to more premenopausal patients being diagnosed with benign conditions (Table 1). Vaginal pH greater than 4.5 typically indicates vaginal dysbiosis, and the frequency of high vaginal pH was 23.15% for benign conditions, 27.78% for hyperplasia, 67.92% for grade 1/2 EEC, and 61.54% for other EC. Vaginal pH significantly differed

between benign and EC groups (p value < 0.0001 and p value 0.006, respectively), as well as hyperplasia and grade 1/2 EEC (p value 0.005). This difference in vaginal pH is unsurprising due to high vaginal pH typically being reported in patients who are postmenopausal and not being treated with hormone replacement therapy³², menopausal status differed between disease groups in our cohort. Other characteristics known to alter the microbiome, such as vaginal douching, tobacco usage, and multiple births, were not significantly different between the groups (Table 1). Our study indicates women diagnosed with EC are more likely to be older, postmenopausal, have a higher BMI, and previously received a D&C. A description of this cohort has been published previously^{33,34}.

Microbiome community structure between patients with EC and with Benign conditions

The composition of the vaginal microbiome in patients with EC was distinct compared to patients with benign conditions using both alpha and beta diversity comparisons). When investigating overall profile differences between malignant and benign conditions at the genus level (Fig. 1A and Supplementary Fig. 1), the vaginal microbiome of patients with EC had significantly higher community richness relative to patients with benign conditions (Faith's PD; Kruskal Wallis; $H = 18.06376$, q -value = 0.00012). There was a significantly increased community richness in postmenopausal patients relative to premenopausal patients (Kruskal Wallis; $H = 11.285466$, q -value = 0.000781), which suggests a depletion of *Lactobacillus* dominance (Supplementary Fig. 1). Since the vaginal microbiome is typically dominated by *Lactobacillus*, a typical genus-level assessment of the vaginal microbiome has low evenness in its microbial diversity, instead, it is dominated by few (or one, in this case) genera³⁵. The microbiome of patients with benign conditions had relatively low evenness, but the patients with EC had significantly increased evenness, again suggesting a shift away from *Lactobacillus* dominance in the vaginal microbiome in patients with EC (Kruskal-Wallis; $H = 18.06376$, q -value = 0.00012). Similarly, there was an increase in evenness when comparing menopausal status (Kruskal-Wallis; $H = 3.920024$, q -value = 0.047714). There was no significant difference in evenness when comparing patients' BMI categories (< 25, 25–29, 30–34, >=35) or grade 1/2 EEC and other ECs. The composition of the vaginal microbiome was also observed to be significantly different between patients with EC and patients with benign conditions, based on Unweighted UniFrac³⁶ distances between samples (PERMANOVA; pseudo- $F = 10.162838$, q -value = 0.0060) (Supplementary Fig. 1).

Since lactobacilli are crucial to vaginal homeostasis, we investigated whether there were differences observed in *Lactobacillus* abundance in vaginal samples. Our analyses revealed there was a significantly lower relative abundance of *Lactobacillus* in patients with EC compared to patients with benign conditions (Mann–Whitney; $U = 4598$, q -value: $3.2e-04$). Aligning with Fig. 1B, this further suggests that patients with EC may have increased microbial diversity. Additionally, patients with grade 3 EEC or other EC subtypes (referred to as other EC) had lower *Lactobacillus* relative frequency compared to grade 1/2 EEC (Mann–Whitney; $U = 438.0$, q -value: $3.087e-02$), possibly indicating that *Lactobacillus* depletion is more profound in severe/aggressive types of cancers. Since previous literature indicates differences amongst vaginal lactobacilli and their association with health³⁷, we investigated lactobacilli differences at the species level (using environment-aware classifiers, which enable increased taxonomic resolution from 16S rRNA reads) but observed no differences between disease groups (Supplementary Fig. 2).

The rectal microbiome composition was significantly different between patients with EC and patients with benign conditions, as tested with Unweighted UniFrac distance (PERMANOVA; pseudo- $F = 4.012069$, q -value = 0.006). However, community richness in the rectal microbiome seems to not differ between patients with benign conditions and patients with EC. Community richness in the rectal microbiome was not significantly different when comparing menopausal status, BMI categories, or EC types.

Table 1 | Demographic/clinical information

	All (n=192)	Benign Gynecologic Conditions (n=108)	Endometrial Hyperplasia (n=18)	Grade 1/2 EEC (n=53)	Other Ecs (n=13)	p value for the difference among diseases		
						overall	Paired	
Age (mean (S.D.)) (n=192)	51.02 (12.45)	45.55 (10.01)	54.11 (13.35)	58.73 (11.82)	60.77 (8.06)	<0.0001	B vs H	0.007
							B vs EEC	<0.0001
							B vs Other EC	<0.0001
Race (n (%)) (n=190)								
American Indian/Alaskan	15 (7.89)	5 (4.67)	1 (5.56)	8 (15.38)	1 (7.69)	0.004	B vs EEC	0.02
Asian/Far East/South East	4 (2.11)	4 (3.74)	0 (0.00)	0 (0.00)	0 (0.00)			
Asian/Indian	1 (0.53)	0 (0.00)	1 (5.56)	0 (0.00)	0 (0.00)			
Native Hawaiian/Pacific Islander	1 (0.53)	0 (0.00)	0 (0.00)	1 (1.92)	0 (0.00)			
White/Caucasian	142 (74.74)	78 (72.90)	16 (88.89)	37 (71.15)	11 (84.62)			
Black or African American	12 (6.32)	11 (10.28)	0 (0.00)	1 (1.92)	0 (0.00)			
Middle Eastern/North African	1 (0.53)	0 (0.00)	0 (0.00)	0 (0.00)	1 (7.69)			
Mixed or Multi-Racial	9 (4.74)	7 (6.54)	0 (0.00)	2 (3.95)	0 (0.00)			
Not specified, other	5 (2.63)	2 (1.87)	0 (0.00)	3 (5.77)	0 (0.00)			
Ethnicity (n=191)								
Non-Hispanic	141 (73.82)	76 (70.37)	14 (77.78)	41 (78.85)	10 (76.92)	0.67		
Hispanic	50 (26.18)	32 (29.63)	4 (22.22)	11 (21.15)	3 (23.08)			
Income (n=168)								
<10,000	8 (4.60)	3 (3.03)	0 (0.00)	5 (11.36)	0 (0.00)			
10,000-25,000	26 (14.94)	10 (10.10)	6 (42.86)	7 (15.91)	3 (27.27)	0.002	B vs H	0.07
25,000-50,000	39 (22.41)	13 (13.13)	3 (21.43)	16 (36.36)	4 (36.36)		B vs EEC	0.0004
50,000-75,000	32 (18.39)	23 (23.23)	2 (14.29)	6 (13.64)	1 (9.09)			
75,000-100,000	18 (10.34)	15 (15.15)	1 (7.14)	1 (2.27)	1 (9.09)			
>100,000	32 (18.39)	24 (24.24)	1 (7.14)	3 (6.82)	2 (18.18)			
Don't know/refused	19 (10.32)	11 (11.11)	1 (7.14)	6 (13.64)	0 (0.00)			
Vaginal pH (n=192)								
≤4.5	118 (61.46)	83 (76.85)	13 (72.22)	17 (32.08)	5 (38.46)			
>4.5	74 (61.54)	25 (23.15)	5 (27.78)	36 (67.92)	8 (61.54)	<0.0001	B vs EEC	<0.0001
							B vs Other EC	0.006
							H vs EEC	0.005
BMI (n=192)								
<25	29 (15.38)	23 (21.30)	0 (0.00)	4 (7.55)	2 (15.38)			
25–29	47 (24.38)	38 (35.19)	1 (5.56)	6 (11.32)	2 (15.38)	<0.0001	B vs H	<0.0001
30–34	30 (15.63)	19 (17.59)	2 (11.11)	6 (11.32)	3 (23.08)		B vs EEC	<0.0001
≥35	86 (44.79)	28 (25.93)	15 (83.33)	37 (69.81)	6 (46.15)			
Menopausal status (n=190)								
Pre	108 (56.84)	89 (82.41)	6 (33.33)	12 (23.53)	1 (7.69)			
Post	82 (43.16)	19 (17.59)	12 (66.67)	39 (76.47)	12 (92.31)	<0.0001	B vs H	<0.0001
							B vs EEC	<0.0001
							B vs Other EC	<0.0001
Combined contraceptives								
Hormonal (n=138)								
Yes	36 (26.09)	26 (32.10)	0 (0.00)	5 (14.71)	5 (45.45)			
No	102 (73.91)	55 (67.90)	12 (100.00)	29 (85.29)	6 (54.55)	0.01	B vs H	0.02
							B vs EEC	0.06
							H vs Other EC	0.01

Table 1 (continued) | Demographic/clinical information

	All (n=192)	Benign Gynecologic Conditions (n=108)	Endometrial Hyperplasia (n=18)	Grade 1/2 EEC (n=53)	Other Ecs (n=13)	p value for the difference among diseases		
						overall	Paired	
Non-hormonal (n=64)							EEC vs Other EC	0.048
Yes	2 (3.13)	1 (2.50)	0 (0.00)	1 (7.14)	0 (0.00)			
No	62 (96.88)	39 (97.50)	6 (100.00)	13 (92.86)	4 (100.00)	0.62		
D&C (n=192)								
Yes	58 (30.21)	19 (17.59)	7 (38.89)	25 (47.17)	7 (53.85)			
No	134 (69.79)	89 (82.41)	11 (61.11)	28 (52.83)	6 (46.15)	0.0002	B vs H	0.056
							B vs EEC	0.0001
							B vs Other EC	0.007
Parity (n=191)								
0	44 (23.04)	22 (20.56)	4 (22.22)	15 (28.30)	3 (23.08)	0.3		
1	24 (12.57)	9 (8.41)	4 (22.22)	8 (15.09)	3 (23.08)			
2	42 (21.99)	23 (21.50)	7 (38.89)	10 (18.87)	2 (15.38)			
3	44 (23.04)	27 (25.23)	2 (11.11)	13 (24.53)	2 (15.38)			
4+	37 (19.37)	26 (24.30)	1 (5.56)	7 (13.21)	3 (23.08)			
Heavy periods (n=150)								
Light	9 (6.00)	5 (5.32)	0 (0.00)	2 (5.88)	2 (20.00)	0.01	B vs EEC	0.01
Mod	48 (32.00)	21 (22.34)	6 (50.00)	17 (50.0)	4 (40.00)		B vs Other EC	0.05
Heavy	93 (62.00)	68 (72.34)	6 (50.00)	15 (44.12)	4 (40.00)			
Tobacco use (n=184)								
Yes	19 (10.33)	14 (13.46)	1 (5.56)	3 (6.12)	1 (7.69)	0.15		
No	55 (29.89)	34 (32.69)	4 (22.22)	13 (26.53)	4 (30.77)			
Never	89 (48.37)	47 (45.19)	12 (66.67)	22 (44.90)	8 (61.54)			
Quit	21 (11.41)	9 (8.65)	1 (5.56)	11 (22.45)	0 (0.00)			
Douching (n=161)								
Yes	26 (16.15)	15 (15.96)	3 (21.43)	7 (16.28)	1 (10.00)	0.9		
No	135 (83.85)	79 (84.04)	11 (78.57)	36 (83.72)	9 (90.00)			
Diabetes (n=192)								
Yes	48 (25.00)	22 (20.37)	5 (27.78)	21 (39.62)	0 (0.00)	0.006	B vs EEC	0.01
No	144 (75.00)	86 (79.63)	13 (72.22)	32 (60.38)	13 (100.00)		H vs Other EC	0.05
							EEC vs Other EC	0.01
Hypertension (n=192)								
Yes	65 (33.85)	25 (23.15)	6 (33.33)	27 (50.94)	7 (53.85)	0.002	B vs EEC	0.001
No	127 (66.15)	83 (76.85)	12 (66.67)	26 (49.06)	6 (46.15)		B vs Other EC	0.04

Values are n (%) unless otherwise stated as mean (SD). P values were calculated using the Kruskal–Wallis test for continuous variables and Fisher's exact test for categorical variables. Significant values ($p < 0.05$) are highlighted in bold. Pairwise comparisons between diagnosis group and abbreviated as mentioned benign conditions (B), hyperplasia (H), grade 1/2 endometroid endometrial carcinoma (EEC), and other endometrial carcinomas (Other EC).

Association of Demographics with Disease Group.

Differentially abundant taxa observed between patients with EC and with Benign conditions

In order to identify microbial signatures that are unique to patients diagnosed with EC, we applied Analysis of Compositions of Microbiomes with Bias Correction (ANCOM-BC)³⁸ (Fig. 2). Vaginal samples were investigated with unadjusted models and revealed three taxa depleted and five increased between EC and benign (Fig. 2A). *Olegusella massiliensis* ($q = 0.004$), *Aerococcus christensenii* ($q = 0.0325$), and *Limosilactobacillus* sp. ($q < 0.0001$) were decreased in EC compared to benign. Vaginal species such as a *Lachnospiraceae* sp. ($q < 0.0001$), *Prevotella timonensis* A ($q = 0.00416$), *Porphyromonas* sp900539765 ($q = 0.0464$), *Peptoniphilus* A sp. ($q = 0.0077$),

and an *Anaerococcus* sp. ($q = 0.00021$) (Fig. 2A, Fig. 2C, Additional File 1) were enriched in EC. Vaginal samples with adjusted models for known risk factors, BMI and menopausal status, revealed 6/8 significant microbes adjusted for BMI only, 4/8 when adjusted for menopausal status, and 2/8 when adjusted for BMI and menopausal status, suggesting that these risk factors may contribute to overall microbial profiles of EC patients (Fig. 2A, Fig. 2C, and Additional File 1). Depletion of *Limosilactobacillus* sp. ($q < 0.0001$) and enrichment of *Lachnospiraceae* sp. ($q < 0.0001$) remained significant for EC compared to benign conditions after adjusting for risk factors. *Veillonella* sp. ($q < 0.0001$), *Fannyhessea vaginae* ($q < 0.0001$), and *Lactobacillus gasseri* ($q < 0.0001$) had a LFC >1e and were significantly

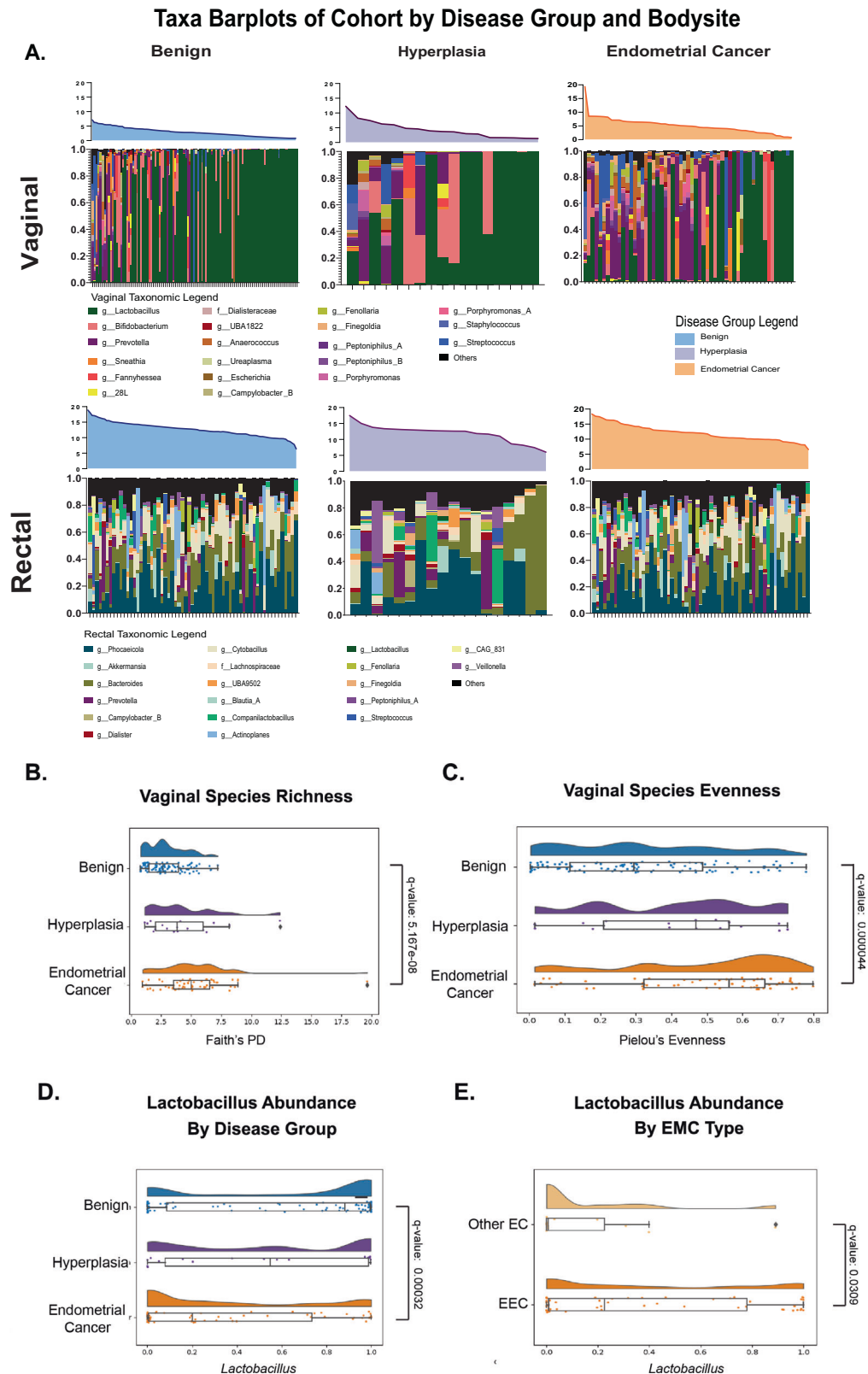
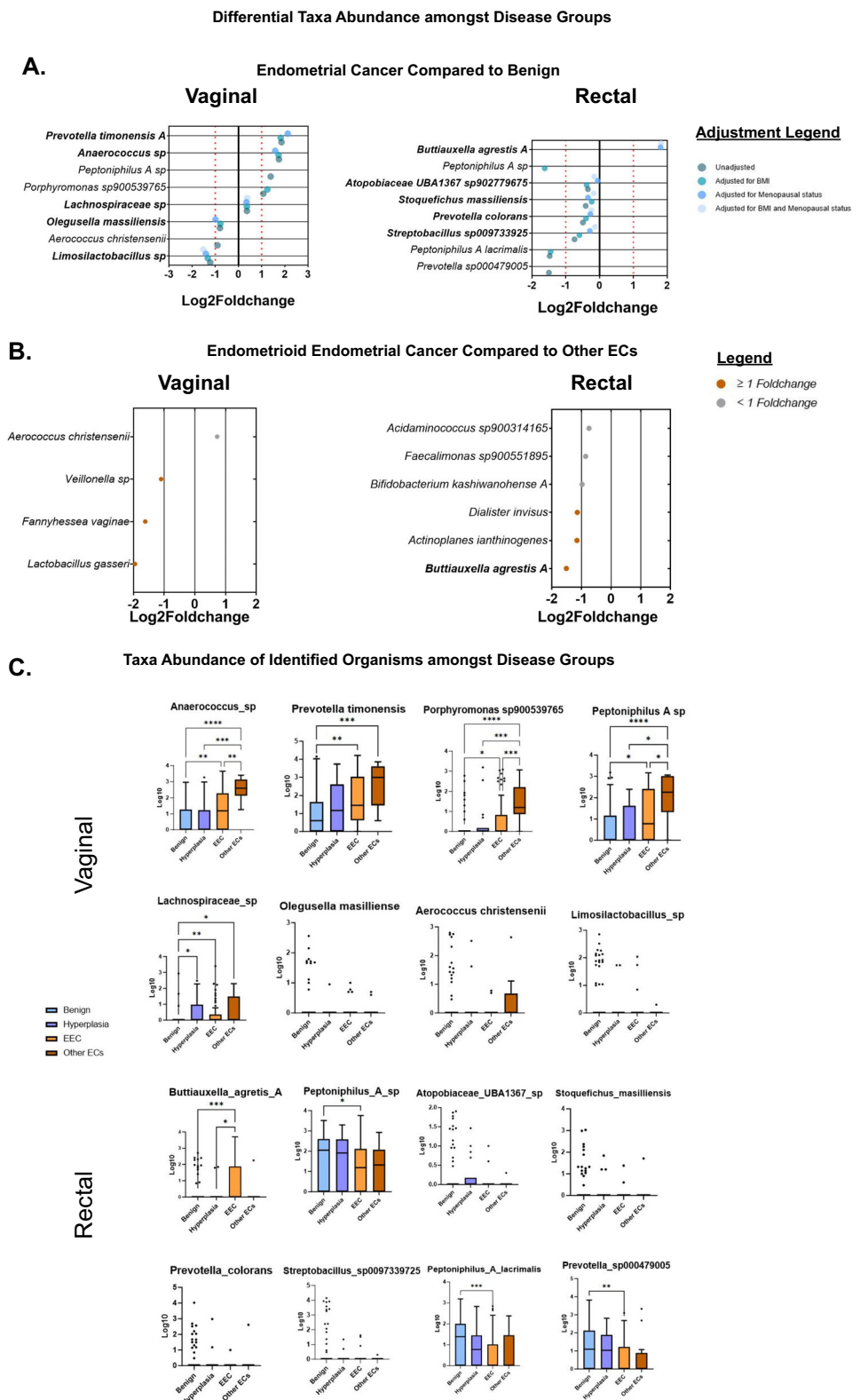


Fig. 1 | Global microbial composition of endometrial cancer patients differs compared to benign gynecologic conditions. The vaginal microbiome is significantly different between patients diagnosed with endometrial cancer and benign conditions. **A** Taxa barplot of genera in the gut and vaginal microbiomes across disease groups: Benign gynecologic conditions (Benign), Endometrial hyperplasia (Hyperplasia), Endometrial cancer. **B** Faith's PD is significantly different between patients with EC and Benign conditions (Kruskal-Wallis: H: 31.784899, p value: 1.722266×10^{-8} , q -value: 5.166797×10^{-8}) and nearly significantly different between

Hyperplasia and Benign individuals (Kruskal-Wallis: H: 3.461674, p value: 6.280670×10^{-2} , q -value: 9.421005×10^{-2}). **C** Pielou's Evenness is also significantly different between EC and benign conditions (Kruskal-Wallis: H: 18.779978, p value: 0.000015 , q -value: 0.000044). **D** There is a significant difference in *Lactobacillus* relative abundance between benign and endometrial cancer individuals (Mann-Whitney-U: q -value: 3.162×10^{-4}). **E** There is also a significant difference between grade 1/2 endometrioid cancer (EEC) and other endometrial cancer subtypes (Other EC) (Mann-Whitney-U: q -value: 3.087×10^{-2}).

Fig. 2 | Dysbiotic microbes in vaginal and rectal microbiomes enriched in endometrial cancer patients compared to benign. *Limosilactobacillus* is depleted, whereas *Anaerococcus*, *Prevotella*, *Porphyromonas*, and *Peptoniphilus* are enriched in EC samples when compared to benign conditions in vaginal samples. *Buttiauxella* is enriched in rectal samples of patients with EC while depleted in *Peptoniphilus* and *Prevotella* in rectal samples. Enrichment in bacterial taxa among EC compared to benign conditions prior (A) and after adjusting for BMI and menopausal status (B). Comparison of differentially abundant taxa of EEC vs. Other ECs with q value of < 0.05 . Dots in dark orange indicate > 1 -Fold change. The bacterial enrichment was performed utilizing ANCOM-BC, visualized where taxa with at least 0.7 LFC and other significant taxa are in the supplement. The threshold for LFC of 1 or greater (orange), below 1 (gray), and all p values were Bonferroni false discovery adjusted. C Boxplots of taxa indicated as significant differentially abundant \log_{10} transformed in this cohort, where the disease group observed abundance. Light blue indicates benign, light purple indicates hyperplasia, light orange indicates EEC and dark orange indicates Other ECs. The Kruskal-Wallis test performed additional abundance differences where * is < 0.05 , ** is < 0.01 , *** < 0.001 , and **** < 0.0001 p value.



depleted in other ECs compared to grade 1/2 EEC (Fig. 2B & Fig. S8 and Additional File 1). Mann-Whitney test revealed that *Lachnospiraceae* sp. ($q = 0.0090$), *Prevotella timonensis* A ($q = 0.0328$), *Porphyromonas* sp900539765 ($q = 0.0002$), *Peptoniphilus* A sp. ($q = 0.0118$), and an *Anaerococcus* sp. ($q = 0.0005$) were enriched in postmenopausal women (Supplementary Fig. 3). *Prevotella timonensis* A ($q = 0.0385$) was only

significantly different when comparing women with a BMI greater than 35 to women with a BMI between 25 and 29 (Supplementary Fig. 4).

Unadjusted models from rectal samples revealed six significantly depleted taxa when comparing EC and benign conditions (Fig. 2A), including *Prevotella* sp000479005 ($q = 0.0076$), *Peptoniphilus* A lacrimalis ($q = 0.0031$), *Streptobacillus* sp009733925 ($q < 0.0001$), *Prevotella colorans*

($q < 0.0001$), *Stoquefichus massiliensis* ($q < 0.0001$), and UBA1367 sp902779675 (from the *Atopobiaceae* family; $q < 0.0001$) (Fig. 2A, C, Additional File 1). Analysis of rectal samples with adjusted models revealed 5/6 significant microbes when adjusted for BMI only, 4/6 when adjusted for menopausal status, and both risk factors (Fig. 2A).

The BMI-adjusted model had an additional taxon not observed in the unadjusted model, a *Peptoniphilus A sp.* ($q < 0.0001$), which was depleted in EC compared to benign conditions. The menopausal status only and menopausal status and BMI adjusted model also included a taxon not observed in the unadjusted model, *Buttiauxella agrestis A* ($q = 0.0205$, $q = 0.0353$), which was enriched in EC compared to benign conditions. Microbes with the log-fold change of greater than one were *B. agrestis A* ($q < 0.0001$), *Bifidobacterium kashiwanohense* ($q < 0.0001$), *Dialister invisus*, and *Actinoplanes ianthogenes* ($q < 0.0001$) were depleted in other ECs compared to grade 1/2 EEC (Fig. 2B, C, S9 Additional File 1). Mann–Whitney test revealed that *Streptobacillus* sp009733925 ($q < 0.0001$) was enriched in premenopausal women, while *Peptoniphilus A lacrimalis* ($q = 0.0059$) was enriched in postmenopausal women (Supplementary Fig. 1A & Supplementary Fig 3). These analyses reveal global and species-specific alterations of vaginal and rectal bacteria in women with EC compared to benign conditions.

Some microbes were differentially abundant in both the vaginal and rectal microbiomes; therefore, we investigated potentially translocated features between body sites to test the hypothesis that the rectal microbiome may be a source (or “reservoir”) of some organisms found in the vaginal microbiome. To test this, we counted the rectal amplicon sequence variants (ASVs) that were also observed in vaginal samples on a per-individual basis. We compared this distribution of counts to what we would expect to see by chance using a bootstrapping test (Mann–Whitney-U one-tailed test $p = 0.0317$)³⁹. Rectal genera observed in both the paired vaginal samples most frequently included *Granulicatella*, *Neisseria*, *Haemophilus D*, *Lactobacillus*, and *Prevotella* (Supplementary Fig 5 & Additional File 2). *Prevotella timonensis* was enriched in EC compared to benign conditions and was in the vaginal samples of 47/90 participants who had *Prevotella timonensis* in their rectal samples. Similarly, *Peptoniphilus A* was enriched in EC vaginal samples compared to benign vaginal samples but was depleted in EC rectal samples compared to benign rectal samples. *Peptoniphilus A* was found in 60/131 participants, with *Peptoniphilus A* in their rectal samples (Additional File 2). Additionally, we observed a significant difference when reversing the directionality of this analysis and counting vaginal microbes that were also observed in rectal microbiomes on a per-individual basis (Mann–Whitney-U one-tailed test $p = 0.0004$). However, the vaginal microbiome as a reservoir only had a median of 20 features that could have been potentially translocated, while the rectal microbiome as a reservoir had a median of 101.5. The presence of an individual’s rectal-associated ASV features in their vaginal microbiome (or vice versa) can represent the transfer of microorganisms between sites. Still, our data cannot confirm that a transfer has occurred as opposed to independent colonization of the two sites by organisms with the same ASV sequence. Putative translocation of rectal to vaginal ASV features among patients with other ECs was higher compared to patients with grade 1/2 EEC (p value = 0.0323) (Supplementary Fig. 6 and Additional File 2). There is a moderately negative correlation between *Lactobacillus* abundance and potential ASV translocation (correlation = -0.425, p value = 5.50e-09). Therefore, the increased potential translocation between the rectal reservoir may indicate dysbiosis in the vaginal microbiome and contribute to gynecologic conditions.

Bacterial community networks differ between patients with EC and Benign conditions

We investigated bacterial co-occurrence networks to understand the relationships between certain bacteria and gain a complete picture of their role in oncogenesis.

Bacterial co-occurrence networks in the rectum of patients with benign conditions had 176 positive and 25 negative correlations (Fig. 3D and

Additional File 3). The genera in the *Peptoniphilaceae* family had many positive correlations ($n = 75$) with microbes associated with gut/rectal health, like *Prevotella* and *Dialister*^{40,41}. Meanwhile, *Lachnospiraceae* family members were found to not co-occur with *Prevotella*. *Lachnospiraceae* have been reported to reduce inflammation in the gut; however, in the vagina, they are associated with bacterial vaginosis⁴². On the other hand, the rectal microbiome of patients with EC had more bacterial community correlations than patients with benign conditions: 263 positive correlations between microbes and 78 negative correlations (Fig. 3C and Additional File 3). We also observed that there were more negative correlations with *Prevotella* and the *Lachnospiraceae* family.

Bacterial co-occurrence networks in the vaginal microbiome also differ between EC and benign. In patients with benign conditions, *Lactobacillus* did not occur with *Prevotella* and *Dialister*, which are associated with bacterial vaginosis (BV). Other BV-associated microbes in the *Peptoniphilaceae* family, such as *Peptoniphilus A*, *Fenollaria*, *Finexgoldia*, and *Anaerococcus*, co-occurred with other BV-associated bacteria: *Porphyromonas* and *Prevotella*, as well as opportunistic pathogens, such as *Campylobacter*. The co-occurrence network of the vaginal microbiome of patients with benign conditions had the smallest network, with 29 positive correlations and two negative correlations (Fig. 3B and Additional File 3). This suggests that the vaginal microbial community structure in patients with benign conditions consists of a smaller group of less diverse microbes compared to patients with EC. Interestingly, *Lactobacillus* was not found to co-occur with any BV-associated microbes, indicating community structures are likely either *Lactobacillus*-dominated or dysbiotic. In comparison, the vaginal microbiome of patients with EC had positive correlations within the *Peptoniphilaceae* family but had more correlations with other BV-associated microbes like *Campylobacter*. The EC group had 43 correlations with the *Peptoniphilaceae* family, and the benign group only had 13. *Lactobacillus* was still negatively correlated with *Prevotella* but also negatively correlated with the 28 L (known as *Megasphaera*) genus (Fig. 3A and Additional File 3). EC also resulted in considerably more correlations between microbes, with 184 positive correlations and 14 negative correlations. Overall, correlations between differentially abundant features suggest that the community networks are different within these microbiomes.

Differences in putative metabolic pathways related to aromatic compounds, antibacterial compounds, and amino acids between EC and benign

To identify potential metabolic pathways driven by bacteria in relation to pathophysiological differences in EC, we utilized PICRUST2⁴³. We observed more putatively altered metabolic pathways from vaginal profiles than rectal profiles prior to adjustment for known risk factors, 86 and two, respectively (Fig. 4A and Supplementary Fig. 7). Vaginal putative metabolic pathways from patients with EC had a depletion in 86 metabolic pathways. Pathways were categorized into subpathways related to amino acid metabolism ($n = 3$), aromatic compound metabolism ($n = 7$), C1 compound utilization and assimilation ($n = 2$), carbohydrate metabolism ($n = 15$), cell structure biosynthesis ($n = 6$), cofactor, carrier, and vitamin biosynthesis ($n = 5$), fatty acid and lipid biosynthesis ($n = 7$), generation of precursor metabolites and energy ($n = 6$), nucleoside and nucleotide biosynthesis ($n = 10$), secondary metabolite biosynthesis ($n = 22$), and other ($n = 3$) (Supplementary Fig. 7 and Additional File 4). Pathways related to antimicrobial biosynthesis with a (LFC) > 1 were (5 R)-carbapenem carboxylate biosynthesis ($q = 0.0019$), butirosin biosynthesis ($q = 0.0014$), clavulanate biosynthesis ($q = 0.0037$), D-cycloserine biosynthesis ($q = 0.0039$), fosfomycin biosynthesis ($q < 0.0001$), neomycin biosynthesis ($q = 0.0054$), paromamine biosynthesis I ($q = 0.0021$), paromamine biosynthesis II ($q = 0.0019$), ribostamycin biosynthesis ($q = 0.0034$), streptomycin biosynthesis ($q < 0.0001$), and superpathway of butirosin biosynthesis ($q = 0.0012$) (Fig. 4A). Interestingly, pathways related to aromatic compound metabolism and lipid metabolism with a LFC > 1 were sitosterol degradation to androstenedione ($q = 0.0033$), benzoyl-CoA degradation II (anaerobic) ($q = 0.0001$), mevalonate pathway I (eukaryotes and bacteria) ($q < 0.0001$) and mevalonate pathway II

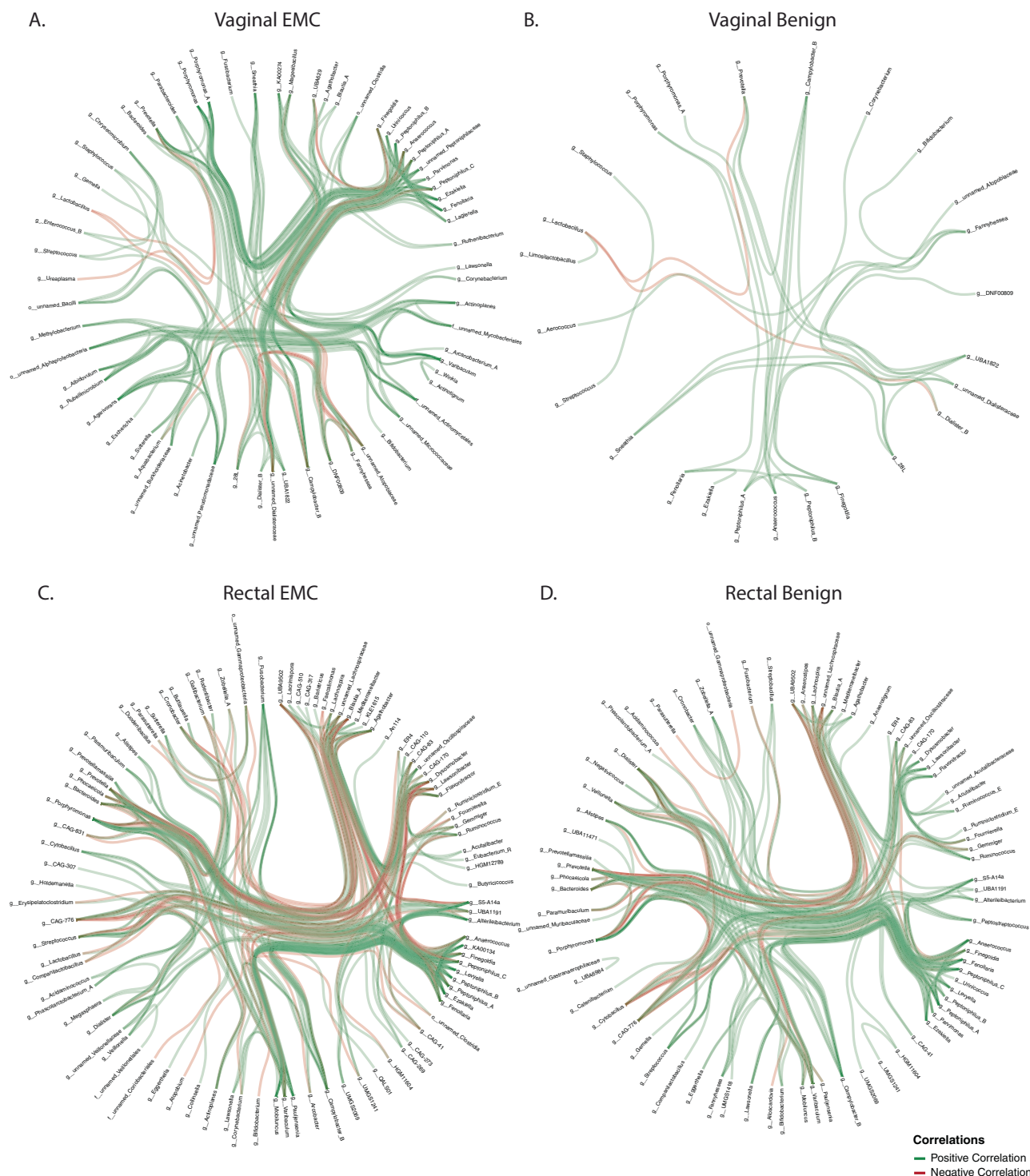


Fig. 3 | Community network differences between EC and benign. The vaginal and rectal microbiome form different community structures when comparing EC communities to benign. **A** The EC vaginal community has more correlations between genera (198 total correlations) than the **(B)** benign vaginal community

(31 total correlations). Similarly, **(C)** the EMC bacterial co-occurrence network of the rectal microbiome (341 total correlations) had more correlations compared to the **(D)** benign community (201 total correlations). Interactive HTML files are featured in the supplement.

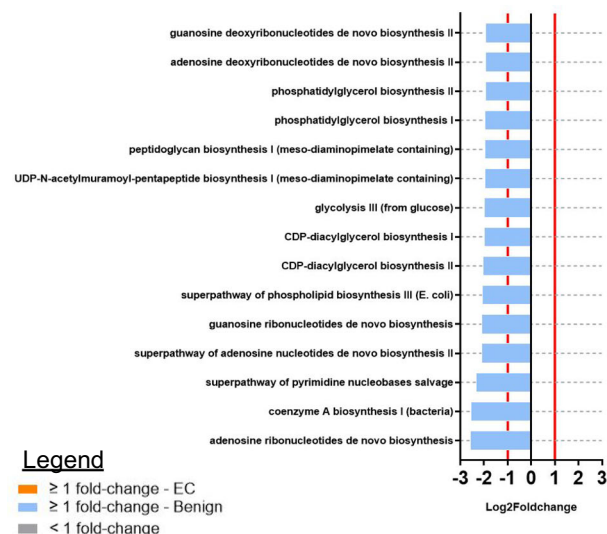
(haloarchaea) ($q = 0.0116$) (Fig. 4A). However, when adjusted for BMI and menopausal status, there was one depleted pathway, L-arabinose degradation IV ($q < 0.0001$), and six enriched pathways: superpathway of L-phenylalanine biosynthesis ($q = 0.0135$), superpathway of L-tyrosine biosynthesis ($q = 0.0171$), glutaryl-CoA degradation ($q = 0.0308$), CMP-3-deoxy-D-manno-octulosonate biosynthesis ($q = 0.0109$), Kdo transfer to lipid IVA ($q = 0.0137$), and lipid IVA biosynthesis ($q = 0.0125$) in patients

with EC (Fig. 4B). Depletion in significant hormone-related pathways was impacted by menopausal status (Supplementary Fig. 7). nly 1,5-anhydrofructose degradation ($q < 0.0001$), part of glycogen degradation, was enriched in grade 1/2 EEC patients compared to other ECs (Fig. 4C).

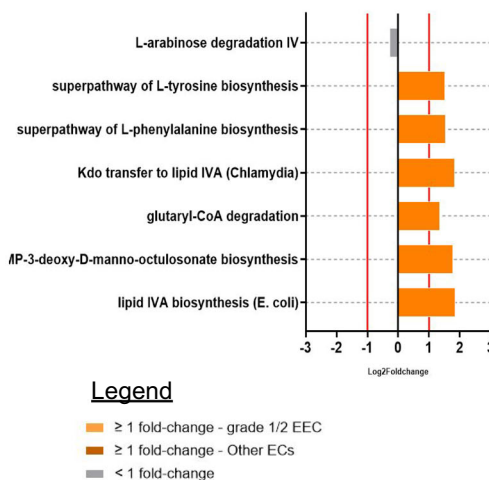
Rectal samples from patients diagnosed with EC had lower putative metabolic contributions of L-valine degradation I ($q = 0.0097$) in unadjusted models and models adjusted for known EC risk factors (Figs. 5A and

Vaginal Depletion of Putative Microbiome-Mediated Hormone Pathways in Endometrial Cancer

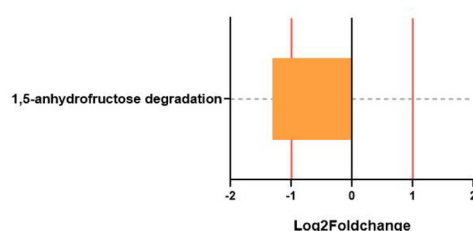
A. Top 15 Pathways Depleted in Endometrial cancer compared to benign (unadjusted)



B. Endometrial cancer compared to benign adjusted for BMI and menopausal status



C. Grade 1&2 EEC compared to Other ECs



D. Significant Putative Pathways Related to Steroid Hormones, Antibiotics, and Carbohydrate Utilization

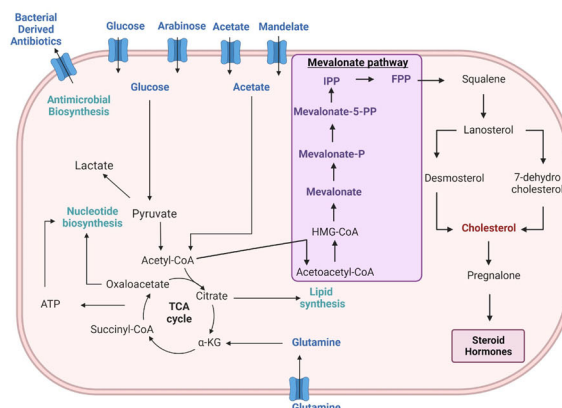


Fig. 4 | Putative vaginal metabolic signatures depleted in endometrial cancer. Cholesterol-dependent pathways are dysregulated in EC samples when compared to benign conditions in vaginal samples. Enrichment in putative metabolic pathways generated by PiCrust2 among EC patients (orange) when compared to benign conditions (light blue) prior (A) and after adjusting for BMI and menopausal status (B). ANCOM-BC was utilized to determine differentially abundant putative pathways where $p < 0.05$ was deemed significant. Pathways with a LFC > 1 are colored

based on disease group; pathways in gray are below LFC < 1.0 . Enrichment in putative metabolic pathways among EC patients, comparing EEC patients (light orange) to other ECs (C). The pathway enrichment was performed utilizing ANCOM-BC. The threshold for log₂ fold difference was 1, and pairwise comparisons were made using Bonferroni multiple testing correction. Pair of pathways relating to cholesterol metabolism that are significantly depleted in EC compared to benign conditions (D).

5B). Unadjusted samples additionally had depletion of dTDP-N-acetylviobiosynthesis ($q = 0.0473$), which is part of O-antigen metabolism and LPS biosynthesis in patients with EC (Fig. 5A, C). Seven pathways were enriched in patients with grade 1/2 EEC compared to other ECs (Fig. 5D, E), most falling into polyphenol degradation pathways such as superpathway of vanillin and vanillate degradation ($q < 0.0001$), vanillin and vanillate degradation I ($q < 0.0001$), vanillin and vanillate degradation II ($q < 0.0001$), syringate degradation ($q < 0.0001$) (Fig. 5D and Fig. 5E). The remaining differential putative metabolic pathways between patients with grade 1/2 EEC and other ECs were super pathway of C_1 compounds oxidation to CO_2 ($q < 0.0001$), L-arabinose degradation IV ($q < 0.0001$), and tylosin biosynthesis ($q < 0.0001$) (Fig. 5D, E). These data further highlight the need to elucidate the role of polyphenol metabolism by microbiota in regard to cancer.

Discussion

Our objective was to evaluate the role of the vaginal and rectal microbiomes as they relate to endometrial cancer¹¹. To address these hypotheses, we evaluated both the vaginal and rectal microbiota profiles using 16S rRNA amplicon sequencing in 192 patients diagnosed with benign conditions ($n = 108$), hyperplasia ($n = 18$), and EC ($n = 66$). Other studies that have investigated the upper female reproductive tract of patients with EC have shown an increase in *Fannyhessea* (formerly *Atopobium vaginae*), *Porphyromonas*, *Anaerococcus*, *Peptoniphilus*, and *Prevotella*^{19,28,29,44}. In Semertzidou et al., *Porphyromonas*, *Prevotella*, *Peptoniphilus*, and *Anaerococcus* were increased in the vagina, cervix, and endometrium³⁰. These studies also observed a depletion in lactobacilli in EC cases compared to benign or healthy control groups^{19,28,30}. Our study identified similar signatures of the vaginal microbiota, such as increased community richness,

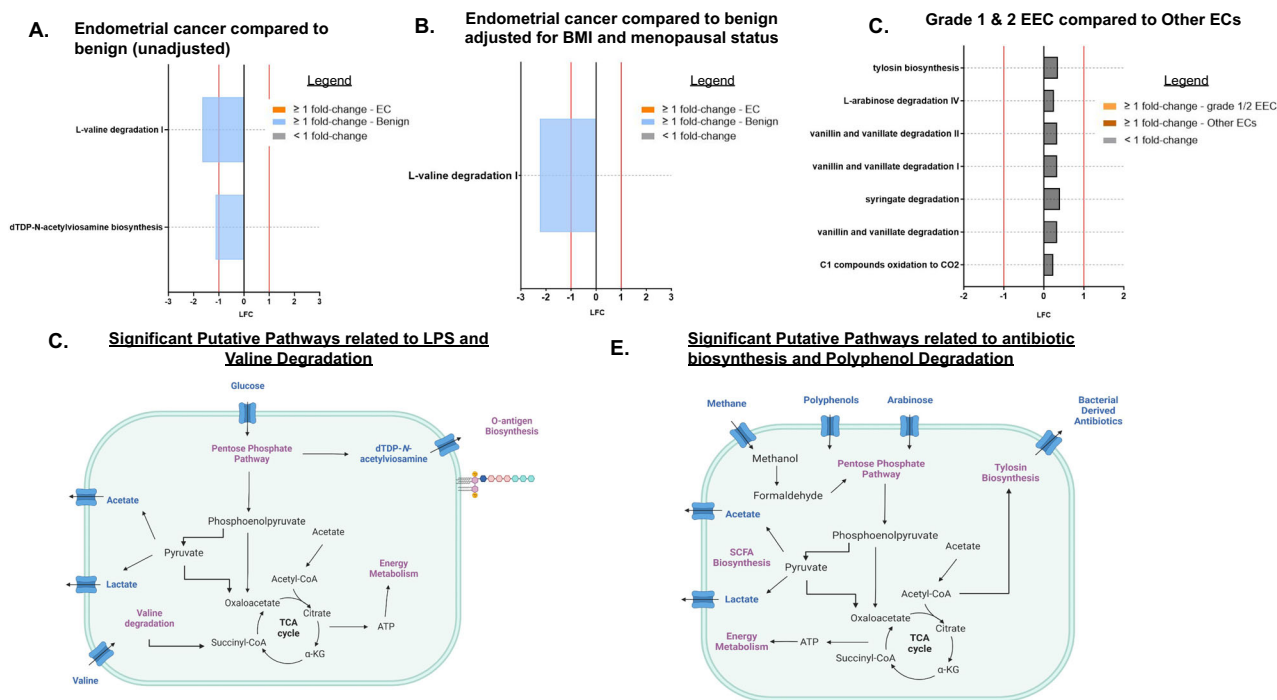


Fig. 5 | Putative metabolic rectal signatures depleted in endometrial cancer.

Amino acid and polyphenol pathways in EC samples when compared to benign conditions in rectal samples. Enrichment in putative metabolic pathways generated by PiCrust2 among EC patients (orange) when compared to benign conditions (light blue) prior (A) and after adjusting for BMI and menopausal status (B). ANCOM-BC was utilized to determine differentially abundant putative pathways where $p < 0.05$ was deemed significant. Pathways with a LFC > 1 are colored based on disease group; pathways in gray are below LFC < 1 . C Map of pathways relating to polyphenol

metabolism and antimicrobial biosynthesis in EC compared to benign conditions. D Enrichment in putative metabolic pathways among EC patients, comparing EEC patients (light orange) to other ECs. The pathway enrichment was performed utilizing ANCOM-BC. The threshold was LFC > 1 , and pairwise comparisons were made using Bonferroni multiple testing correction. E Map of pathways relating to polyphenol metabolism and antimicrobial biosynthesis in 1/2 EEC compared to Other ECs.

depletion of lactobacilli, and differentially abundant taxa such as *Anaerococcus*, *Prevotella*, *Porphyromonas*, and *Peptoniphilus* in patients with EC compared to benign gynecologic conditions (Fig. 6). These recurring signatures/patterns in multiple reports and sites support the hypothesis of the ascension of vaginal microbes to endometrium^{45–47}, where they can interact with endometrial epithelial cells and impact carcinogenesis or features of the tumor microenvironment^{48,49}. Further, we identified vaginal microbes such as *Veillonella*, *F. vaginae*, and *L. gasseri* to be differentially enriched in grade 1/2 EEC. In contrast, other more aggressive EC subtypes were enriched in *A. christensenii*, which could be of interest in differentiating types of EC. In addition, our study revealed that in EC, vaginal networks of microbiota increased in the number of positive microbe-microbe correlations, specifically within the family *Peptoniphilaceae*, suggesting microbe-microbe interactions associated with cancer (Fig. 6). Further investigation into these bacteria, their interactions, and their role in inflammation could provide key insights for the development and etiology of EC¹¹.

Additionally, the gut microbiomes' contribution to estrogen-dependent gynecologic conditions such as EC is hypothesized to occur through estrogen metabolism mediated by beta-glucuronidases of gut bacteria, termed the estrobolome^{17,50}. Furthermore, it has been observed that the gut microbiome is altered during obesity⁵¹, a known risk factor for EC. Obesity has been linked to estrogen dysregulation as adipose tissue can increase circulating estrogen levels, especially in postmenopausal women^{17,52,53}. Our study identified one microbe in the rectal microbiome, *B. agretis*, after adjusting for menopausal status and BMI enriched in EC. This is an under-described gram-negative *Enterobacteriaceae* species but has been found in clinical specimens of wounds^{54–56}. Future investigation into this opportunistic microorganism is needed but may hold insight into inflammatory bacteria associated with gastrointestinal and gynecologic cancers¹¹. Li et al. observed Proteobacteria, Gammaproteobacteria,

Enterobacteriales, *Enterobacteriaceae*, and *Shigella* in the gut of patients with EC compared to healthy controls²⁷. Further, our study observed species in the rectal microbiome that differentiated EC subtypes, including *B. kashiwanohense*, *D. invisus*, and *A. ianthogenes*, which were most enriched in grade 1/2 EEC compared to other EC (Fig. 6). Interestingly, *Bifidobacterium* has been shown to metabolize estrogen in the gut, which could be why it is linked with grade 1/2 EEC, as these are estrogen-mediated ECs⁵⁷.

EC has known risk factors, such as BMI and menopausal status^{58,59}, and in this cohort, the mean BMI for grade 1/2 EEC was 40.29, and for other ECs, it was 37.22. Among patients with grade 1/2 EEC, 76.47% were postmenopausal, while 92.31% of patients with other ECs were postmenopausal. Since these factors outside of EC have been linked to microbial changes in vaginal and gut/rectal microbiomes^{60–62}, herein we account for these factors in microbial associations. Vaginal *Lachnospiraceae* sp., *P. timonensis* A, *Porphyromonas* sp900539765, *Peptoniphilus* A sp., and *Anaerococcus* sp. were enriched in postmenopausal women. However, when adjusting for menopause status alone, *Lachnospiraceae* sp., *P. timonensis* A, and *Anaerococcus* sp. remained enriched, and *O. massiliense* (a reclassified microbe from the *Atopobiaceae* family) and a *Limosilactobacillus* sp. remained depleted in EC. Only a *Lachnospiraceae* sp. and a *Limosilactobacillus* sp. remained significant after adjustment for BMI and menopause status. Walsh et al. noted that menopausal status was the greatest factor to impact the microbiome of women with EC, with *Anaerococcus*, *Porphyromonas*, and *Peptoniphilus* being the most predictive of menopausal status in their study²⁹. Since these microbes have had associations with EC in other studies, it is likely due to a strong link with postmenopausal women diagnosed with EC. Only *P. lacrimalis* was associated with postmenopausal status in the rectal microbiome samples in our study. Further investigation is needed into *Peptoniphilus* species from both mucosal sites related to EC and postmenopausal status.

Microbiome in Endometrial Cancer Compared to Benign Gynecologic Conditions

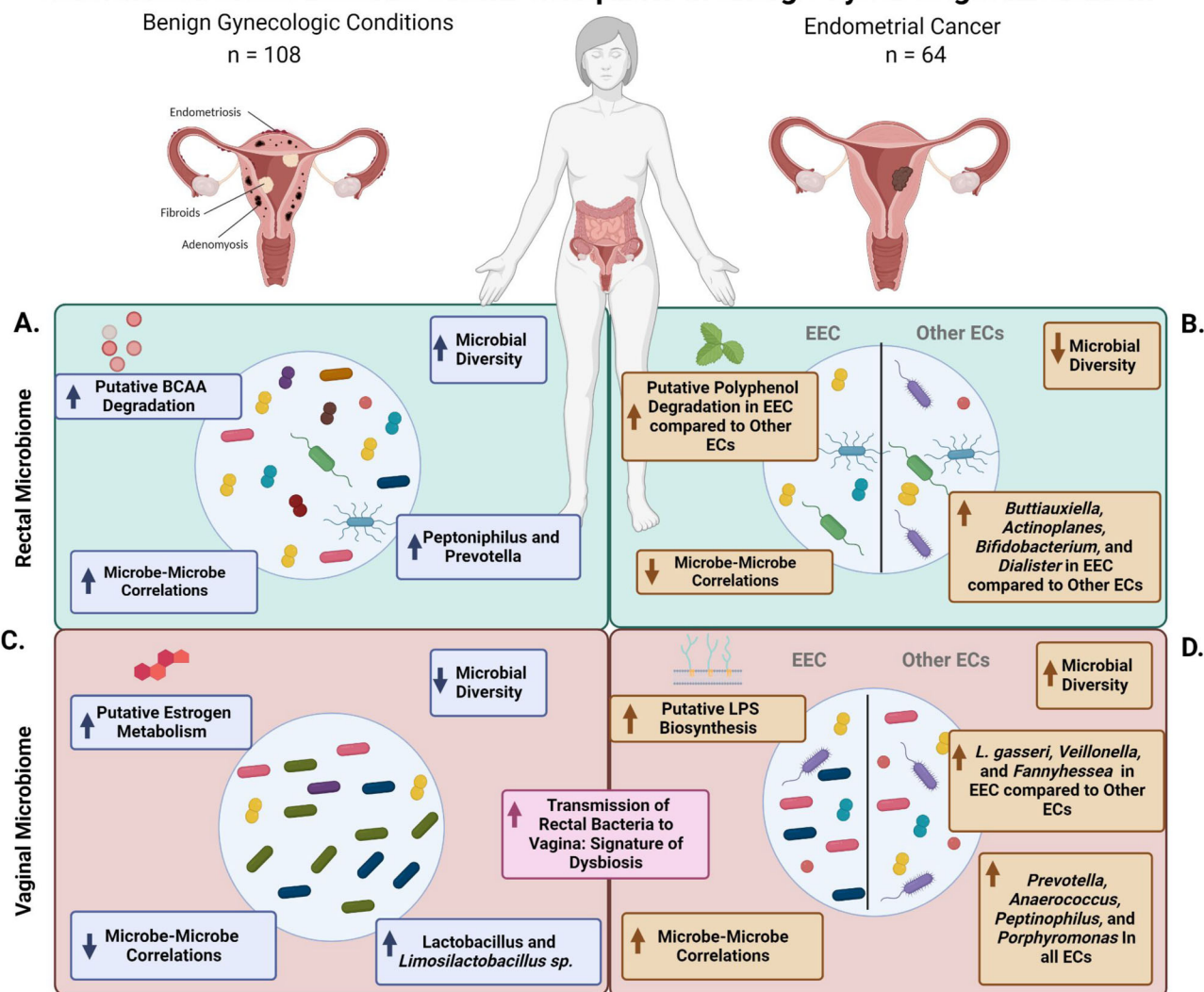


Fig. 6 | Summary of microbiome findings in endometrial cancer compared to benign gynecologic conditions. Diagram of major findings between endometrial cancer and benign gynecologic conditions observed by vaginal and rectal microbiome. **A** The rectal microbiome of patients with benign gynecologic conditions such as endometriosis, adenomyosis, and fibroids was revealed to have increased microbial diversity, increased abundance of *Peptoniphilus* and *Prevotella* species, increased bacterial networks, and increased valine degradation pathways when compared to patients with endometrial cancer. **B** The rectal microbiome of patients with the diagnosis of endometrial cancer, either low-grade endometrioid carcinoma or other endometrial cancer subtypes, revealed to have an increase in the abundance of *Buttiauxiella agretis*, less bacterial community networks and an increased potential translocation of microbes from the rectum to vagina. **C** The vaginal microbiome of

patients with benign gynecologic conditions was revealed to have increased *Lactobacillus* abundance and, specifically, an increased abundance of a particular *Limosilactobacillus* species; most bacterial networks were negatively correlated. In addition, multiple putative metabolic pathways relating to amino acid metabolism, steroid metabolism, and antibiotic metabolism were enriched when compared to profiles from patients with endometrial cancer. **D** The vaginal microbiome of patients with EC was revealed to have increased microbial diversity as well as increased *Prevotella*, *Anaerococcus*, *Peptoniphilus*, and *Porphyromonas* abundance. Most bacterial networks were positively correlated with *Prevotella*, indicating the highest correlations. In addition, *Lactobacillus gasseri*, *Veillonella* sp., and *Fannyhessea vaginae* were increased in low-grade endometrioid carcinoma compared to other endometrial cancer subtypes.

Our results indicate, based on microbial feature sharing analysis, that as the vaginal microbiome shifts towards dysbiosis, more gut microbes become increasingly present in the vaginal microbiome, with the most transfer of *Prevotella timonensis* and *Peptoniphilus* A. Notably, in our study, we also observed both *Prevotella timonensis* and *Peptoniphilus* to be of higher abundance in EC. One study hypothesized that a shorter anovaginal distance might increase the microbial sharing from the rectum to the vagina. Still, they found no significant associations with anovaginal distance and increased risk of BV, even though transmission of microbes did occur⁶³. Our findings suggest that the rectal microbiome might be an extrvaginal reservoir for a dysbiotic vaginal microbiome and importantly that an increased abundance of vaginal *Lactobacillus* may be protective against the transfer of rectal microbes (Fig. 6). More aggressive subtypes of EC did have

the most microbial feature sharing. To our knowledge, there are no studies that have investigated the vaginal microbiome as a reservoir for gut/rectal microbes in gynecologic cancers. Our study observed that shared features occurred bidirectionally, which may signal overall dysbiosis in gynecologic conditions. The mechanism of potential translocation in either direction is unknown, and future studies may further reveal microbial translocation in vivo and its association with vaginal and gut health. We hypothesize that *Prevotella timonensis* may be one of the most transferred microbes between vaginal and rectal mucosal sites due to its sialidase activity that degrades mucin glycans or complex carbohydrates, thereby deteriorating the protective mucosal barriers and promoting translocation^{64,65}. Further, this capability of using an alternative food source of mucin may provide a way for *P. timonensis*

to outcompete other microorganisms and relate to its pathogenicity and role in numerous gynecologic conditions⁶⁶.

The transition from carbohydrate to amino acid metabolism has been associated with BV and cancer cell survival and proliferation^{67,68}. In accordance with this, we found that in vaginal samples, there were 15 pathways depleted in carbohydrate metabolism. Previous literature has also indicated that carbohydrates are increased in profiles with a high abundance of *Lactobacillus*⁶⁹; thus, this metabolic signature may be driven by the *Lactobacillus* depletion observed in EC. Further, a specific pathway involved in glycogen degradation, 1,5-anhydrofructose degradation, was enriched in patients with grade 1/2 EEC compared to other EC subtypes (Fig. 6). Glycogen is an important complex carbohydrate in the vaginal microenvironment associated with *Lactobacillus* dominance, and glycogen synthesis is stimulated by estrogen⁷⁰. This supports the *Lactobacillus* differences observed in grade 1/2 EEC compared to other ECs. Our study observed lipid metabolism pathways related to steroid hormones in EC, including sitosterol degradation to androstenedione, benzoyl-CoA degradation II, and mevalonate metabolism. A previous study investigated uterine tumor microbial profiles of early-stage EC between White and Black women and identified microbial pathways in mevalonate to be significantly different between Black and White women diagnosed with EC⁷¹. These pathways were lost after adjustment for BMI and menopausal status, which is unsurprising due to the link between hormone dysregulation and vaginal microbiota alterations observed in women who have a higher BMI or who are postmenopausal^{26,60,72}. Further, our study identified that putative metabolic contributions of bacteria were more observed in vaginal samples ($n = 86$) than in rectal samples ($n = 2$) when comparing EC.

Putative pathway analysis of rectal samples also revealed that patients with EC have depleted valine degradation (Fig. 6). Some studies suggest the role of bacteria in modulating branched-chain amino acids in the gut, like valine. This could make amino acids available to the tumor microenvironment for host cell activation of mTORC1 signaling, which can increase protein translation and cell growth⁷³. Interestingly, three vanillic acid degradation and syringate degradation pathways, which are all polyphenol degradation pathways, were differentially increased in other EC subtypes (Fig. 6). Polyphenols are commonly found in food sources and have been proposed as antioxidants with anti-cancer effects^{74,75}; thus, degradation of these polyphenols by microbiota may decrease these beneficial effects. Additionally, polyphenols and estradiol both have phenolic structures, thus polyphenols may have the ability to impact on E2 and estrogen receptor signaling⁷⁶ which could be an important factor for estrogen-dependent cancers such as grade 1/2 EEC. Thus, could be why degradation of these phenolic compounds is enriched in Other EC subtypes which are typically estrogen independent. Further research into these metabolic pathways in the vaginal and gut/rectal environments is still needed to confirm these findings.

Other strengths of the study include strict inclusion/exclusion criteria to account for external factors that impact the microbiome, such as antibiotic usage and histopathological confirmation of diagnoses used to stratify patients into disease groups. Further, sampling was collected at time of hysterectomy and consisted of women with newly diagnosed EC prior to treatment. We conducted an analysis of two mucosal body sites, the vagina and the rectum, for disease comparisons. We implemented new bioinformatics tools to investigate bacterial feature transfer between these mucosal sites. We also adjusted our analyses for multiple risk factors of EC, known as modulators of the human microbiome. We acknowledge that this study is utilizing 16S rRNA as a bacterial marker gene, and therefore, the full taxonomic resolution and metabolic potential for species identification may be limited. A small sample size of the endometrial hyperplasia and racial/ethnic groups prevented us from conducting some analyses due to the unequal sample distribution between groups. Thus, future studies should include more patients with endometrial hyperplasia and patients from more diverse backgrounds to better identify the role of the microbiome in EC development and whether the microbiome plays

a role in EC health outcomes such as prognosis and mortality between different racial/ethnic communities. It is important to note that these findings compare patients diagnosed with EC and patients with benign gynecologic conditions, two populations that undergo hysterectomy. Therefore, future studies may assess non-surgical populations to assess key differences between populations that do not present with symptoms or undergo gynecologic surgeries and those with diagnosed pathologies. That said, there are limitations to comparing these populations, including a lack of histological confirmation of disease that may indicate underdiagnosed benign conditions. Additionally, investigating environmental factors such as phthalate exposure, medication usage, and diet could be of interest to further refine the complexities of the microbiome in cancer.

This study provides foundational knowledge of the global microbial alterations from multiple mucosal body sites in EC patients, further revealing associations of polymorphic microbiomes within endometrial cancer. This finding is supported by our analyses identifying increased community richness and significant differences in community composition in vaginal and rectal microbial profiles of patients diagnosed with EC. Interestingly, vaginal and rectal microbiome alterations were observed in the overall community network structures, with more microbe-microbe correlations, specifically in the family *Peptoniphilaceae*, observed in patients diagnosed with EC. Further, unique microbial features of EC, such as a depletion of *Lactobacillus* and an increase of *Anaerococcus*, *Porphyromonas*, and *Peptoniphilus* in vaginal samples, and a depletion of *Peptoniphilus* and *Prevotella* and an increase in *Buttiauxella* in rectal samples were observed. Transfer of microbiota between mucosal sites was also observed, with *Prevotella timonensis* and *Peptoniphilus A. sp.* being most shared between sites, and the transfer was reduced as *Lactobacillus* dominance increased. We also identified putative metabolic pathways that link microbiota to pathophysiology, with enrichment of L-tyrosine and L-phenylalanine biosynthesis in EC relating to dysregulation of amino acids, enrichment of 1,5 anhydrofructose degradation relating to complex carbohydrate degradation in grade 1/2 EEC, and depletion in mevalonate metabolism and enrichment of cholesterol metabolism relating to hormone metabolism in patients with EC. These findings set the framework for future microbiome and metabolomics integration studies, using whole metagenomic sequencing to identify microbial capabilities for estrogen metabolism and investigating other mucosal sites, such as the endometrium. Further research on these topics could identify critical microbes and pathways associated with EC, potentially leading to novel approaches for preventing and treating EC using microbiome modulation. This study provides foundational knowledge on the role of polymorphic microbiomes in EC and key microbial targets for future investigation to assess host-microbe interactions related to the etiology of disease.

Methods

Patient recruitment

We recruited 192 non-pregnant women who were undergoing hysterectomy. These women were grouped by the following conditions based on histopathological diagnosis: grade 1/2 endometrial endometrioid carcinoma (EEC) ($n = 53$), other EC subtypes ($n = 13$), endometrial hyperplasia ($n = 18$), and benign conditions ($n = 108$). Benign conditions often co-occur, and our study population consists of patients diagnosed with endometriosis ($n = 21$), fibroids ($n = 70$), adenomyosis ($n = 46$), or other gynecologic conditions ($n = 18$) that have a histopathological diagnosis, such as abnormal uterine bleeding (Fig. 7). These patients were recruited between June 2018 and February 2020 from three clinical sites in the Phoenix (AZ, USA) metropolitan area: Banner University Medical Center – Phoenix, Valleywise Health Medical Center, and Dignity Health Chandler Regional Medical Center. The Institutional Review Board of the University of Arizona approved this study (reference no. 1708726047), and written informed consent was obtained from each study participant. The study was conducted in accordance with federal guidelines, regulations, and the Declaration of Helsinki. We included women of any race or ethnicity aged 18 years or older.

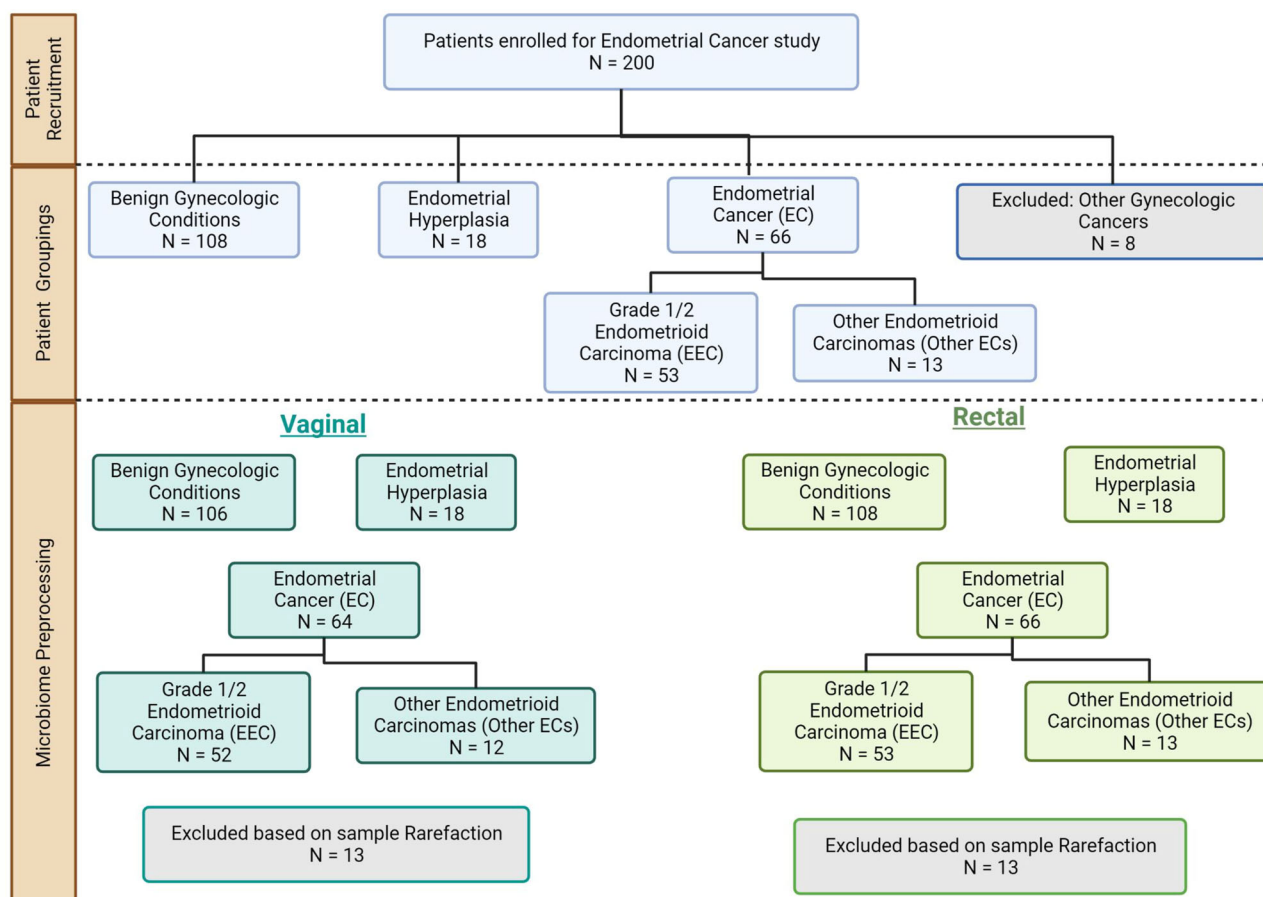


Fig. 7 | Study and sample enrollment diagram. $N = 200$ women enrolled in the Endometrial Cancer study with histologically confirmed diagnoses of benign gynecological conditions ($n = 108$), endometrial hyperplasia ($n = 18$), grade 1/2 endometrioid endometrioid carcinoma ($n = 56$), other endometrioid carcinomas

($n = 13$), and other gynecologic cancers. $n = 8$ women were diagnosed with other gynecologic cancers, so were excluded from the analysis, resulting in our study population of $n = 192$ women. The number of samples excluded after microbiome preprocessing were $n = 13$ for vaginal and $n = 13$ for rectal.

Patients were excluded based on the following criteria: if they were currently menstruating; currently lactating; currently or within the past three months using antibiotics, antifungals, or antivirals; or currently or within the past three weeks had a vaginal infection (including bacterial vaginosis and candidiasis), vulvar infection, urinary tract infection, or sexually transmitted infection (including chlamydia, gonorrhea, trichomoniasis, genital herpes, HIV). Patients were also excluded if they used vaginally applied medications and suppositories, feminine deodorant sprays, wipes, or vaginal lubricants within 48 h of the visit; used depilatory treatments in the genital area within 72 h of the visit; or had sex within 48 h of the visit. Lastly, individuals were excluded if they went bathing or swimming within the previous 4 h, consumed nicotine products within the previous 2 h, or had a diagnosis of diabetes, hepatitis, or any condition in the genital area interfering with the study. The exclusion criteria were verified by physician pelvic exam, medical record, and/or self-report. Demographic, socioeconomic, and medical history data were collected from surveys and/or medical records. Further information on inclusion and exclusion criteria and detailed clinical and demographic information were described previously³³.

Sample collection and processing

Clinical samples were collected by a surgeon prior to hysterectomy and surgical sterilization. A speculum was inserted without lubricant, and two vaginal swabs were collected. The first vaginal swab was collected by swabbing the lateral walls of the mid vagina using an eSwab collection system containing Amies transport medium (COPAN Diagnostics, Murrieta, CA). The second vaginal swab was used to measure vaginal pH using nitrazine paper and recorded by the clinician

according to the manufacturer's instructions using a scale of 4.5–7.5. Following vaginal swab collection, two FLOQ swabs (COPAN Diagnostics, Murrieta, CA) were individually inserted into the rectum, swabbing inner rectal tissue. The specimen swabs were immediately placed on ice and frozen at -80°C within one hour of collection. On the day of extraction, swabs were thawed on ice, and DNA was extracted using a DNeasy PowerSoil Isolation Kit (Qiagen, Germantown, MD) following the manufacturer's instructions. Sample DNA was aliquoted to avoid freeze/thaw cycles and stored at -80°C for further analysis.

DNA sequencing

The quality of the extracted DNA samples was checked using Qubit Fluorometers (ThermoFisher). The 16S rRNA amplicon libraries were constructed using the Earth Microbiome Project (EMP) primers (515F-806R) with Golay barcode tags on the forward primer⁷⁷ (PCR mix ratio is shown in Supplementary Table 1). The annealing was performed for 20 s at 50°C . The 16S rRNA fragment was extended for 45 s at 72°C . The PCR for both vaginal and rectal samples used 30 cycles. The quality of amplicons was checked via agarose gel electrophoresis. The amplicon library of the samples that did not have a clear DNA band, or those showing DNA bands in the negative control lane of the agarose gel electrophoresis, were reconstructed until they passed this quality check. The concentration of the quality-assured amplicons was measured using a fluorescent-based DNA measuring technique, Quant-iT dsDNA High-Sensitivity Assay (Invitrogen). 200 nM of each amplicon was pooled into two library pools for sequencing. The quality of the pool was verified

with the Bioanalyzer DNA 1000 chip (Agilent Technologies, Santa Clara, CA), and then the pool was combined with 1% PhiX for sequencing. The pools were sequenced on the Illumina MiSeq using the 600-cycle MiSeq reagent kit version 3 (Illumina, San Diego, CA) using the Illumina MiSeq benchtop sequencing platform.

Demultiplexing

The raw data obtained for this study was multiplexed paired-end sequences in Earth Microbiome Project protocol format⁷⁷. Demultiplexing was performed using QIIME 2's `qiime demux emp-paired`^{78,79} command.

Human contamination filtration

Human contamination filtering was performed in two steps. First, reads that matched the human genome were removed from the data. Then, post-quality-control amplicon sequence variants (ASVs) that were not assigned to a bacterial phylum were removed. This two-step process should filter all human sequences from a 16S rRNA amplicon data set. These two steps are described in detail here: The demultiplexed data was exported from QIIME 2 and aligned to a hg19 reference human⁸⁰ genome using bowtie2. The SAM file output provided by Bowtie was converted to a BAM file using samtools⁸¹ and then converted to a FASTQ file using Bedtools⁸². Before aligning to the reference human genome, there were 13,152,450 sequences in the vaginal and rectal microbiome data combined. After aligning, there were 13,145,634 sequences retained; less than 0.06% of the sequences were filtered out due to human contamination. This methodology is detailed in the QIIME 2 Cancer Microbiome Intervention tutorial at <https://docs.qiime2.org>.

After performing quality control (see below), amplicon sequence variants that were not taxonomically classified to at least the phylum level were filtered out using `qiime feature-table filter-features`. In the gut microbiome data, there were originally 3624 ASVs, and 16% of the ASVs were filtered out (3046 remaining ASVs). In the vaginal microbiome data, there were originally 1386 ASVs, and 13.3% of the ASVs were filtered out (1201 remaining ASVs).

Quality control

Quality control was performed on the sequences by running `'qiime dada2 denoise-paired'`⁸³. The forward sequences were trimmed at 240, and the reverse sequences were trimmed at 220 due to reduced sequence quality beginning at these positions. After quality control, the resulting feature table was split into two groups based on the body site from which the sample was collected (vaginal or rectal) using `qiime feature-table filter-samples`.

Taxonomic classification

The rectal and vaginal microbiomes were taxonomically classified using environment-weighted classifiers. The weights tailor the classifier towards the microbiome that is being classified and enable higher-resolution taxonomic assignment without the need for an environment-specific database (which can prevent uncommon taxa from being reported). Both classifiers were built using the 202 release of Genome Taxonomy Database (GTDB)^{84–87} using `'qiime feature-classifier fit-classifier-naive-bayes'`. For the rectal microbiome, `'qiime clawback generate-class-weights'`^{88,89}, based on publicly available stool data in Qiita⁹⁰, was used to generate environment-specific taxon weights. `'qiime feature-classifier fit-classifier-naive-bayes'` combined the class weights with the base classifier. The classifier was used to assign taxonomic labels to ASVs using `'qiime feature-classifier classify-sklearn'`. The vaginal classifier was created using methods outlined in Bokulich et al. 2022⁹¹ to create vaginal weights based on the stirrup database⁹². Then `'qiime feature-classifier classify-sklearn'` was run to annotate the vaginal microbiome ASVs taxonomically.

Microbiome diversity analysis

Diversity analyses were performed separately on each body site. A rooted phylogenetic tree was created using `'qiime phylogeny align-to-tree-mafft-`

`fasttree'`⁹³. `'qiime diversity core-metrics-phylogenetic'` was run to compute diversity metrics. The sampling depth was set at 15,993 for the rectal microbiome and 3714 for the vaginal microbiome. The sampling depths were confirmed to be representative of the microbiome's diversity using `qiime diversity alpha-rarefaction`. `'qiime diversity alpha-group-significance'` was used to compare Faith's phylogenetic diversity (PD)⁹⁴ and evenness. Of the 192 total samples, 13 vaginal samples and 13 separate rectal samples were excluded after microbiome rarefaction (Fig. 7). Statistical comparisons were made using Kruskal-Wallis⁹⁵, and all comparisons were corrected using the false discovery rate (FDR) method⁹⁶. `'qiime diversity beta-group-significance'` was used to generate Unweighted Unifrac distances⁹⁷ between samples, and statistical comparisons were made using PERMANOVA.

Lactobacillus abundance analysis

To investigate *Lactobacillus* abundance, ASVs were collapsed to the genus or species level using `'qiime taxa collapse'`. The collapsed feature tables were then transformed into relative frequency using `'qiime feature-table relative-frequency'`. The relative frequencies were then compared using Scipy stat's Mann-Whitney-U⁹⁸ statistical tests, and the comparisons were corrected using statsmodels FDR corrections⁹⁶. This analysis was visualized using Seaborn⁹⁹ and Plotly¹⁰⁰ to create raincloud plots.

Differentially abundant taxa analysis

To identify differentially abundant taxa in rectal or vaginal samples between EC and benign groups, as well as between grade 1/2 EEC and other EC, taxa counts at the species-level were run through the package Analysis of Compositions of Microbiomes with Bias Correction (ANCOM-BC)³⁸ utilizing R version 4.2.2¹⁰¹. To address risk factors associated with EC, such as menopausal status (pre- and post-menopausal) and BMI (< 25, 25–29, 30–34, and ≥35), additional adjustment analyses were performed. Comparisons of the EC to benign groups were adjusted for menopausal status alone, BMI alone, and menopausal status and BMI together. *P* values were corrected for multiple comparisons using the FDR method, and the term “*q* value” is used to refer to these FDR-corrected *p* values in this work. Taxa with a *q*-value < 0.05 were considered significant. Taxa with a (LFC) > 1 were considered biologically differential, and dot plot visualizations of these data were created in Prism 9.0 (GraphPad, San Diego, CA).

Bacterial correlation

Using the genus table, `'qiime SCNIC sparcc-filter'`^{102,103} was applied to remove features observed fewer than 500 times and features present in less than two samples. `'qiime SCNIC calculate-correlations'`¹⁰⁴ was run with default settings to generate correlations between microbes in a microbiome. These correlations were calculated for each body site's benign and EC groups. These commands were run using QIIME 2 2020.10 because of incompatibilities between the q2-scnic QIIME 2 plugin and QIIME 2 2022.2, which was used for the remainder of the analyses here.

Rectal reservoir analysis

q2-fmt (<https://github.com/qiime2/q2-fmt>) was used to test the co-occurrence of ASVs between the rectal microbiome and the vaginal microbiome, which we use here as a proxy for potential transfer from the rectal to the vaginal microbiome. The rectal microbiome was considered the “donor,” and the vaginal microbiome was considered the “recipient”. `'qiime fmt sample-peds'`³⁹ was run to calculate a proportion of rectal features that “transferred” to the vaginal microbiome. We additionally tested this using the vaginal microbiome as the “donor” and the rectal microbiome as the “recipient”. To assess whether a given number of “transferred” microbiomes was more than should be expected by chance, the per-subject vaginal-rectal pairings were randomized 999 times to generate “randomized Proportional Engraftment of Donor Strains (PEDS) values.” The code for the bootstrapping verification is Additional Data 5. The real vaginal-rectal pairs' PEDS values were compared to randomized PEDS values with a one-tailed Mann-Whitney-U test⁹⁸. This tests if vaginal and rectal microbiomes from the same individual had more potential translocated features than

microbiomes from randomly associated vaginal-rectal pairings, providing an assessment of how likely a PEDS score is to be observed if there is no potential for transfer between an individual's vaginal and rectal samples¹⁰⁵. PEDS values were correlated with *Lactobacillus* relative frequency using Spearman's Correlation.

Putative metabolic contributions of microbiome analysis

To determine the putative metabolic contributions from vaginal and rectal taxa identified in this cohort, QIIME 2¹⁰⁶ was utilized to run Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUST2)⁴³ using the q2-picrust2 QIIME 2 plugin on 16S rRNA read counts. Host and microbial genomes from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database¹⁰⁷ and the PICRUST2 predicted metagenomes were used to predict the functional gene content of the 16S rRNA profiles. Pathways were labeled based on MetaCyc¹⁰⁸ pathways. Statistical analysis was performed utilizing ANCOM-BC³⁸ to determine differentially abundant putative metabolic pathways between EC and benign groups and between grade 1/2 EEC and other EC. To control for risk factors associated with EC, such as menopausal status (pre- and post-menopausal) and BMI (< 25, 25–29, 30–34, and ≥35), additional adjustment analyses were performed on the EC compared to benign groups for menopausal status and BMI together. *P* values were corrected for multiple comparisons using the FDR method, and taxa with a *q* value < 0.05 were considered significant. Putative metabolic pathways with a LFC > 1 were considered biologically differential, and diverging bar charts were visualized in Prism 9.0 (GraphPad, San Diego, CA).

Data availability

16S rRNA gene sequences were deposited in the National Center for Biotechnology Information (NCBI) database on BioProject accession number PRJNA1036657.

Code availability

A bioinformatics reproducibility supplement¹⁰⁹, providing full detail on all QIIME 2 steps of this analysis, is provided as Additional File 6.

Received: 19 July 2024; Accepted: 1 April 2025;

Published online: 22 May 2025

References

- Bray, F. et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* [Internet]. **68**, 394–424 (2018).
- Zhang, S. et al. Global, regional, and national burden of endometrial cancer, 1990–2017: Results from the global burden of disease study, 2017. *Front Oncol.* [Internet]. **9**, 1440 (2019).
- Rahib, L. et al. Projecting cancer incidence and deaths to 2030: The unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res.* [Internet]. **74**, 2913–2921 (2014).
- Chatterjee, S., Gupta, D., Caputo, T. A. & Holcomb, K. Disparities in gynecological malignancies. *Front Oncol.* [Internet]. **6**, 36 (2016).
- Stewart, C. E., Nañez, A., Ayoola-Adeola, M. & Chase, D. Reducing health disparities in endometrial cancer care in 2024. *Curr. Opin. Obstet. Gynecol.* [Internet]. **36**, 18–22 (2024).
- Mayo Clinic [Internet]. [cited 2023 Sep 1]. Endometrial cancer. Available from: <https://www.mayoclinic.org/diseases-conditions/endometrial-cancer/symptoms-causes/syc-20352461> 2023
- Makker, V. et al. Endometrial cancer. *Nat. Rev. Dis. Prim.* [Internet]. **7**, 88 (2021).
- Cancer Genome Atlas Research Network, Kandoth, C. et al. Integrated genomic characterization of endometrial carcinoma. *Nat.* [Internet]. **497**, 67–73 (2013).
- Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: The next generation. *Cell* [Internet]. **144**, 646–674 (2011).
- Hanahan, D. Hallmarks of cancer: New dimensions. *Cancer Discov.* [Internet]. **12**, 31–46 (2022).
- Łaniewski, P., İlhan, Z. E. & Herbst-Kralovetz, M. M. The microbiome and gynaecological cancer development, prevention and therapy. *Nat. Rev. Urol.* [Internet]. **17**, 232–250 (2020).
- Shahanavaj, K. et al. Cancer and the microbiome: potential applications as new tumor biomarker. *Expert Rev. Anticancer Ther.* [Internet]. **15**, 317–330 (2015).
- Nomura, A. et al. Helicobacter pylori infection and gastric carcinoma among Japanese Americans in Hawaii. *N. Engl. J. Med.* [Internet]. **325**, 1132–1136 (1991).
- Kostic, A. D. et al. Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. *Genome Res* [Internet]. **22**, 292–298 (2012).
- Ahn, J. et al. Human gut microbiome and risk for colorectal cancer. *J. Natl. Cancer Inst.* [Internet]. **105**, 1907–1911 (2013).
- Łaniewski, P. et al. Linking cervicovaginal immune signatures, HPV and microbiota composition in cervical carcinogenesis in non-Hispanic and Hispanic women. *Sci. Rep.* [Internet]. **8**, 7593 (2018).
- Boutriqu S. et al. Gut and endometrial microbiome dysbiosis: A new emergent risk factor for endometrial cancer. *J. Pers. Med.* [Internet]. **11**, 659 (2021).
- Shanmughapriya, S. et al. Viral and bacterial aetiologies of epithelial ovarian cancer. *Eur. J. Clin. Microbiol Infect. Dis.* [Internet]. **31**, 2311–2317 (2012). SepAvailable from.
- Walther-Antônio, M. R. S. et al. Potential contribution of the uterine microbiome in the development of endometrial cancer. *Genome Med* [Internet]. **8**, 122 (2016). Nov 25Available from.
- Banerjee, S. et al. The ovarian cancer oncobiome. *Oncotarget* [Internet]. **8**, 36225–36245 (2017).
- Martin, D. H. & Marrazzo, J. M. The vaginal microbiome: Current understanding and future directions. *J. Infect. Dis.* [Internet]. **214**, S36–S41 (2016).
- Gajer, P. et al. Temporal dynamics of the human vaginal microbiota. *Sci. Transl. Med* [Internet]. **4**, 132ra52 (2012).
- Mirmonsef, P. et al. Glycogen levels in undiluted genital fluid and their relationship to vaginal pH, estrogen, and progesterone. *PLoS One* [Internet]. **11**, e0153553 (2016).
- Takada, K. et al. Female reproductive tract-organ axes. *Front Immunol.* [Internet]. **14**, 1110001 (2023).
- Qi, X., Yun, C., Pang, Y. & Qiao, J. The impact of the gut microbiota on the reproductive and metabolic endocrine system. *Gut Microbes* [Internet]. **13**, 1–21 (2021).
- Graham, M. E. et al. Gut and vaginal microbiomes on steroids: implications for women's health. *Trends Endocrinol. Metab.* [Internet]. **32**, 554–565 (2021).
- Li, Y., Liu, G., Gong, R. & Xi, Y. Gut microbiome dysbiosis in patients with endometrial cancer vs. healthy controls based on 16S rRNA gene sequencing. *Curr. Microbiol* [Internet]. **80**, 239 (2023).
- Gressel, G. M., Usyk, M., Frimer, M., Kuo, D. Y. S. & Burk, R. D. Characterization of the endometrial, cervicovaginal and anorectal microbiota in post-menopausal women with endometrioid and serous endometrial cancers. *PLoS One* [Internet]. **16**, e0259188 (2021).
- Walsh, D. M. et al. Postmenopause as a key factor in the composition of the Endometrial Cancer Microbiome (ECbiome). *Sci. Rep.* [Internet]. **9**, 19213 (2019).
- Semertzidou A. et al. Microbial signatures and continuum in endometrial cancer and benign patients [Internet]. Research Square. Available from: <https://www.researchsquare.com/article/rs-2102199/v1> 2022
- Crosbie, E. J. et al. Endometrial cancer. *Lancet* [Internet]. **399**, 1412–1428 (2022). Apr 9Available from.
- Panda, S., Das, A., Singh, A. S. & Pala, S. Vaginal pH: A marker for menopause. *J. Midlife Health* [Internet]. **5**, 34–37 <https://pmc.ncbi.nlm.nih.gov/articles/PMC3955044/> (2014).

33. Łaniewski, P. et al. Protein biomarkers in cervicovaginal lavages for detection of endometrial cancer. *Biomark. Res [Internet]*. **10**, 88 (2022).
34. Lorentzen G. M. et al. Cervicovaginal metabolome and tumor characteristics for endometrial cancer detection and risk stratification. *Clin. Cancer Res.* **30**, 3073–3087 (2024).
35. Miller, E. A., Beasley, D. E., Dunn, R. R. & Archie, E. A. Lactobacilli dominance and vaginal pH: Why is the human vaginal microbiome unique?. *Front Microbiol [Internet]*. **7**, 1936 (2016).
36. Hamady, M., Lozupone, C. & Knight, R. Fast UniFrac: Facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *ISME J. [Internet]*. **4**, 17–27 (2010). JanAvailable from.
37. Chee, W. J. Y., Chew, S. Y. & Than, L. T. L. Vaginal microbiota and the potential of Lactobacillus derivatives in maintaining vaginal health. *Micro Cell Fact. [Internet]*. **19**, 203 (2020). Nov 7Available from.
38. Lin, H. & Peddada, S. D. Analysis of compositions of microbiomes with bias correction. *Nat. Commun. [Internet]*. **11**, 3514 (2020).
39. Aggarwala, V. et al. Precise quantification of bacterial strains after fecal microbiota transplantation delineates long-term engraftment and explains outcomes. *Nat. Microbiol [Internet]*. **6**, 1309–1318 (2021).
40. Precup, G. & Vodnar, D. C. Gut Prevotella as a possible biomarker of diet and its eubiotic versus dysbiotic roles: A comprehensive literature review. *Br. J. Nutr. [Internet]*. **122**, 131–140 (2019).
41. Lal S., Kandiyal B., Ahuja V., Takeda K., Das B. Chapter Seven - Gut microbiome dysbiosis in inflammatory bowel disease. In: Das B, Singh V, editors. Progress in Molecular Biology and Translational Science [Internet]. Academic Press; p. 179–204. Available from: <https://www.sciencedirect.com/science/article/pii/S1877117322001326> 2022.
42. Vacca, M. et al. The controversial role of human gut lachnospiraceae. *Microorganisms*. **8**, 573 <https://doi.org/10.3390/microorganisms8040573> (2020).
43. Douglas, G. M. et al. PICRUST2 for prediction of metagenome functions. *Nat. Biotechnol. [Internet]*. **38**, 685–688 (2020).
44. Hakimjavadi, H. et al. The vaginal microbiome is associated with endometrial cancer grade and histology. *Cancer Res. Commun. [Internet]*. **2**, 447–455 (2022). JunAvailable from.
45. Quayle, A. J. The innate and early immune response to pathogen challenge in the female genital tract and the pivotal role of epithelial cells. *J. Reprod. Immunol. [Internet]*. **57**, 61–79 (2002).
46. Hansen, L. K. et al. The cervical mucus plug inhibits, but does not block, the passage of ascending bacteria from the vagina during pregnancy. *Acta Obstet. Gynecol. Scand. [Internet]*. **93**, 102–108 (2014).
47. Audirac-Chalifour, A. et al. Cervical microbiome and cytokine profile at various stages of cervical cancer: A pilot study. *PLoS One [Internet]*. **11**, e0153274 (2016).
48. Crooks T. A., et al. Porphyromonas somerae invasion of endometrial cancer cells. *Front Microbiol [Internet]*. 2021 Jul 23 [cited 2023 Oct 13];12. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8343132/>
49. Caselli, E. et al. Atopobium vaginae and Porphyromonas somerae induce proinflammatory cytokines expression in endometrial cells: A possible implication for endometrial cancer?. *Cancer Manag. Res. [Internet]*. **11**, 8571–8575 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6767476/> (2019).
50. Salliss, M. E., Farland, L. V., Mahnert, N. D. & Herbst-Kralovetz, M. M. The role of gut and genital microbiota and the estrobolome in endometriosis, infertility and chronic pelvic pain. *Hum. Reprod. Update [Internet]*. **28**, 92–131 (2021).
51. Liu, B. N., Liu, X. T., Liang, Z. H. & Wang, J. H. Gut microbiota in obesity. *World J. Gastroenterol. [Internet]*. **27**, 3837–3850 (2021).
52. Acharya, K. D. et al. Distinct changes in gut microbiota are associated with estradiol-mediated protection from diet-induced obesity in female mice. *Metabolites [Internet]*. **11**, 499 (2021).
53. Schreurs, M. P. H., de Vos van Steenwijk, P. J., Romano, A., Dieleman, S. & Werner, H. M. J. How the gut microbiome links to menopause and obesity, with possible implications for endometrial cancer development. *J. Clin. Med. Res. [Internet]* **10**, 2916 (2021).
54. Nakamichi, N., Moriuchi, R., Dohra, H., Futamata, H. & Tashiro, Y. Complete genome sequence of butiauxella agrestis DSM 9389. *Microbiol Resour. Announc [Internet]*. **10**, e00301–e00321 (2021).
55. Antonello, V. S. et al. FB. Post-cesarean surgical site infection due to Butiauxella agrestis. *Int J. Infect. Dis. [Internet]*. **22**, 65–66 (2014).
56. Patra, N., Prakash, M. R., Patil, S. & Rao, M. B. First case report of surgical site infection due to butiauxella agrestis in a neurocare center in India. *Arch. Med. Health Sci. [Internet]*. **6**, 117 https://journals.lww.com/armh/Fulltext/2018/06010/First_Case_Report_of_Surgical_Site_Infection_Due.18.aspx (2018).
57. Flores, R. et al. Fecal microbial determinants of fecal and systemic estrogens and estrogen metabolites: a cross-sectional study. *J. Transl. Med [Internet]*. **10**, 253 (2012). Dec 21Available from.
58. Jo, H. et al. Metabolic syndrome as a risk factor of endometrial cancer: A nationwide population-based cohort study of 2.8 million women in South Korea. *Front Oncol. [Internet]*. **12**, 872995 (2022).
59. Tzenios N., Chahine M., Tazani M. Obesity and endometrial cancer: the role insulin resistance and adipokines. *ISSN 2976-5609 [Internet]*. [cited 2023 Oct 7];1(2). Available from: <https://sjmas.com/index.php/sjmas/article/view/12>
60. Łaniewski, P. & Herbst-Kralovetz, M. M. Connecting microbiome and menopause for healthy ageing. *Nat. Microbiol [Internet]*. **7**, 354–358 (2022).
61. Park M. G., Cho S., Oh M. M. Menopausal changes in the microbiome-A review focused on the genitourinary microbiome. *Diagnostics (Basel) [Internet]*. 2023;13 1193.
62. Garg, A. et al. Vaginal microbiome in obesity and its impact on reproduction. *Best. Pr. Res Clin. Obstet. Gynaecol. [Internet]*. **90**, 102365 (2023).
63. Brotman, R. M., Melendez, J. H. & Ghanem, K. G. A case control study of anovaginal distance and bacterial vaginosis. *Int J. STD AIDS [Internet]*. **22**, 231–233 (2011).
64. İlhan, Z. E., Łaniewski, P., Tonachio, A. & Herbst-Kralovetz, M. M. Members of Prevotella genus distinctively modulate innate immune and barrier functions in a human three-dimensional endometrial epithelial cell model. *J. Infect. Dis. [Internet]*. **222**, 2082–2092 (2020).
65. Segui-Perez C. et al. Prevotella timonensis degrades the vaginal epithelial glycocalyx through high fucosidase and sialidase activities [Internet]. bioRxiv. 2024 [cited 2024 Apr 17]. p. 2024.01.09.574844. Available from: <https://www.biorxiv.org/content/10.1101/2024.01.09.574844v1>
66. Pelayo P., et al. Prevotella are major contributors of sialidases in the human vaginal microbiome [Internet]. bioRxiv. 2024 [cited 2024 Apr 17]. p. 2 Available from: <https://www.biorxiv.org/content/10.1101/2024.01.09.574895v1>
67. Srinivasan, S. et al. Metabolic signatures of bacterial vaginosis. *MBio [Internet]*. **6**, e00204–e00215 (2015).
68. Wei, Z., Liu, X., Cheng, C., Yu, W. & Yi, P. Metabolism of amino acids in cancer. *Front Cell Dev. Biol. [Internet]*. **8**, 603837 (2020).
69. Microbiome Research Reports [Internet]. [cited 2023 Sep 26]. Available from: <https://www.oaepublish.com/mrr/article/view/4861>
70. Mirmonsef, P. et al. Free glycogen in vaginal fluids is associated with Lactobacillus colonization and low vaginal pH. *PLoS One [Internet]*. **9**, e102467 (2014).
71. Hawkins, G. M. et al. Differences in the microbial profiles of early stage endometrial cancers between Black and White women. *Gynecol. Oncol. [Internet]*. **165**, 248–256 (2022).
72. Allen, N. G. et al. The vaginal microbiome in women of reproductive age with healthy weight versus overweight/obesity. *Obes. [Internet]*. **30**, 142–152 (2022).

73. Jung, M. K., Okekunle, A. P., Lee, J. E., Sung, M. K. & Lim, Y. J. Role of branched-chain amino acid metabolism in tumor development and progression. *J. Cancer Prev. [Internet]*. **26**, 237–243 (2021).
74. Rani, R. et al. A landscape analysis of the potential role of polyphenols for the treatment of Polycystic Ovarian Syndrome (PCOS). *Phytomedicine [Internet]*. **2**, 100161 <https://www.sciencedirect.com/science/article/pii/S2667031321001433> (2022).
75. Kaur, J., et al. Discovering multifaceted role of vanillic acid beyond flavours: Nutraceutical and therapeutic potential. *Trends Food Sci. Technol. [Internet]*. **122**, 187–200 <https://www.sciencedirect.com/science/article/pii/S092422442200007> (2022).
76. Cipolletti, M., Solar Fernandez, V., Montalesi, E., Marino, M. & Fiocchetti, M. Beyond the antioxidant activity of dietary polyphenols in cancer: The modulation of estrogen receptors (ERs) signaling. *Int J. Mol. Sci. [Internet]*. **19**, 2624 <https://pmc.ncbi.nlm.nih.gov/articles/PMC6165334/> (2018).
77. Caporaso, J. G. et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J. [Internet]*. **6**, 1621–1624 (2012).
78. Hamady, M., Walker, J. J., Harris, J. K., Gold, N. J. & Knight, R. Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat. Methods [Internet]*. **5**, 235–237 (2008).
79. Hamady, M. & Knight, R. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Res. [Internet]*. **19**, 1141–1152 (2009).
80. Church, D. M. et al. Modernizing reference genome assemblies. *PLoS Biol. [Internet]*. **9**, e1001091 (2011).
81. Danecek P., et al. Twelve years of SAM tools and BCF tools. *Gigascience* **10**, giab008. <https://doi.org/10.1093/gigascience/giab008> (2021).
82. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinforma. [Internet]*. **26**, 841–842 (2010).
83. Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods [Internet]*. **13**, 581–583 (2016).
84. Parks, D. H. et al. A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat. Biotechnol. [Internet]*. **38**, 1079–1086 (2020).
85. Parks, D. H. et al. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res [Internet]*. **50**, D785–D794 (2022).
86. Rinke, C. et al. A standardized archaeal taxonomy for the Genome Taxonomy Database. *Nat. Microbiol [Internet]*. **6**, 946–959 (2021).
87. Parks, D. H. et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol. [Internet]*. **36**, 996–1004 (2018).
88. Bokulich, N. A. et al. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome [Internet]*. **6**, 90 (2018).
89. Kaehler, B. D. et al. Species abundance information improves sequence taxonomy classification accuracy. *Nat. Commun. [Internet]*. **10**, 4643 (2019).
90. Gonzalez, A. et al. Qiita: rapid, web-enabled microbiome meta-analysis. *Nat. Methods [Internet]*. **15**, 796–798 (2018).
91. Bokulich, N. A. et al. Multi-omics data integration reveals metabolome as the top predictor of the cervicovaginal microenvironment. *PLoS Comput Biol. [Internet]*. **18**, e1009876 (2022).
92. Fettweis, J. M. et al. Species-level classification of the vaginal microbiome. *BMC Genomics [Internet]*. **13**, S17 (2012).
93. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2-approximately maximum-likelihood trees for large alignments. *PLoS One [Internet]*. **5**, e9490 (2010). Mar 10 Available from.
94. Faith, D. P. Conservation evaluation and phylogenetic diversity. *Biol. Conserv [Internet]*. **61**, 1–10 <https://www.sciencedirect.com/science/article/pii/S0006320792912013> (1992).
95. Kruskal, W. H. & Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc. [Internet]*. **47**, 583–621 <https://www.tandfonline.com/doi/abs/10.1080/01621459.1952.10483441> (1952).
96. Benjamini Y., Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Series B Stat Methodol [Internet]*. 1995;57:289–300. Available from: <http://www.jstor.org/stable/2346101>
97. Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J. & Knight, R. UniFrac: an effective distance metric for microbial community comparison. *ISME J. [Internet]*. **5**, 169–172 (2011).
98. Mann, H. B. & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *aoms [Internet]*. **18**, 50–60 <http://projecteuclid.org/euclid.aoms/1177730491> (1947).
99. Waskom, M. seaborn: statistical data visualization. *J. Open Source Softw. [Internet]*. **6**, 3021 <https://joss.theoj.org/papers/10.21105/joss.03021> (2021).
100. Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R. & Kievit, R. A. Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Res. [Internet]*. **4**, 63 (2019).
101. Team R. C., Team R. C., Others. RA language and environment for statistical computing. R Foundation for Statistical Computing, Vienna-References. Scientific Research Publishing; 2020.
102. Hagberg A., Swart P. J., Schult D. A. Exploring network structure, dynamics, and function using NetworkX [Internet]. Los Alamos National Laboratory (LANL), Los Alamos, NM (United States); 2008 Jan [cited 2023 Aug 18]. Report No.: LA-UR-08-05495; LA-UR-08-5495. Available from: <https://www.osti.gov/biblio/960616>
103. Watts S. C., Ritchie S. C., Inouye M., Holt K. E. FastSpar: Rapid and scalable correlation estimation for compositional data [Internet]. bioRxiv. 2018 [cited 2023 Aug 18]. p. 272583. Available from: <https://www.biorxiv.org/content/early/2018/03/23/272583>
104. Friedman, J. & Alm, E. J. Inferring correlation networks from genomic survey data. *PLoS Comput Biol. [Internet]*. **8**, e1002687 (2012).
105. Herman C., Gehret L., Bolyen E., Gregory Caporaso J. q2-FMT [Internet]. 2023. Available from: <https://zenodo.org/record/8364982>
106. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol. [Internet]*. **37**, 852–857 (2019). Aug Available from.
107. Kanehisa, M. The KEGG database. *Novartis Found. Symp. [Internet]*. **247**, 91–101 <https://www.ncbi.nlm.nih.gov/pubmed/12539951> (2002).
108. Caspi, R. et al. The MetaCyc database of metabolic pathways and enzymes - a 2019 update. *Nucleic Acids Res. [Internet]*. **48**, D445–D453 (2020).
109. Keefe C. R. et al. Facilitating bioinformatics reproducibility [Internet]. arXiv [q-bio.QM]. Available from: <http://arxiv.org/abs/2305.11198> 2023.

Acknowledgements

We would like to thank the patients who enrolled in the study. This study was primarily supported by the Mary Kay Foundation Translational Research Grant to M.M.H-K. and the Valleywide Research Partnership Grant (VRP26) to M.M.H-K and D.M.C. This work was also funded in part by NIH NCI Informatics Technology for Cancer Research Award 1U24CA248454-01 to J.G.C. N.R.J. was supported by the Sidney Hopkins, Mayola B. Vail, and Patricia Ann Hanson Postdoctoral Fellowship award from the Community Foundation for Southern Arizona. N.R.J. and P.L. are part of the Guiding U54 Investigator Development to Sustainability (GUIDeS) shared resource through the Partnership for Native American Cancer Prevention; the partnership is funded under parallel grants, U54CA143924 and U54CA143925 and are supporting early-stage investigators. We would like to thank the following individuals for their assistance in this study as follows. Regina Montero and Elisa Martinez for their assistance in patient recruitment,

sample, and clinical data collection. Dr. Denise Roe who assisted in statistical consultation of the analyses. Dr. Evan Bolyen and Colin V. Wood for their technical support in bioinformatics. Kathryn A. Conn and Emily M. Borsom for their assistance in prepping, amplifying, and sequencing microbial DNA. Finally, Mary Jewell and Carol Haussler for additional grammar and flow feedback for the final publication. This study was primarily supported by the Mary Kay Foundation Translational Research Grant to M.M.H.-K. and the Valleywide Research Partnership Grant (VRP26) to M.M.H.-K. and D.M.C. This work was also funded in part by NIH NCI Informatics Technology for Cancer Research Award 1U24CA248454-01 to J.G.C. N.R.J. was supported by the Sidney Hopkins, Mayola B. Vail, and Patricia Ann Hanson Post-doctoral Fellowship award from the Community Foundation for Southern Arizona. N.R.J. and P.L. are part of the Guiding U54 Investigator Development to Sustainability (GUIDeS) shared resource through the Partnership for Native American Cancer Prevention; the partnership is funded under parallel grants, U54CA143924 and U54CA143925 and are supporting early-stage investigators.

Author contributions

M.M.H.-K. and D.M.C. conceived this study, constructed the experimental design, and were responsible for funding acquisition. M.M.H.-K. led the overall direction, planning, and coordination of the multi-site study and provided supervision for the project. J.G.C. provided supervision of microbiome analyses and funding acquisition. J.M., N.D.M., and D.C. participated in patient recruitment, sample, and data collection. P.L. assisted in sample processing and DNA extractions. E.C. and K.L. assisted in library preparation and performed sequencing for the study. N.R.J. and C.R.H. conducted all analyses, drafted visualizations, and wrote the first draft of the manuscript. N.R.J., C.R.H., P.L., J.G.C. and M.M.H.-K. contributed to the interpretation of the results and data visualization. N.R.J., C.R.H., P.L., E.C., K.L., J.M., N.D.M., and D.C., J.G.C. and M.M.H.-K. edited, revised, and approved the final manuscript.

Competing interests

M.M.H.-K. serves on the scientific advisory board for Freya Biosciences. None of this work related to, was shared with or was licensed to this

company or any other commercial entity. N.R.J., C.R.H., P.L., E.C., K.L., J.M., N.D.M., D.C., and J.G.C. declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41522-025-00690-1>.

Correspondence and requests for materials should be addressed to J. Gregory Caporaso or Melissa M. Herbst-Kralovetz.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025