SUPPLEMENTAL MATERIALS

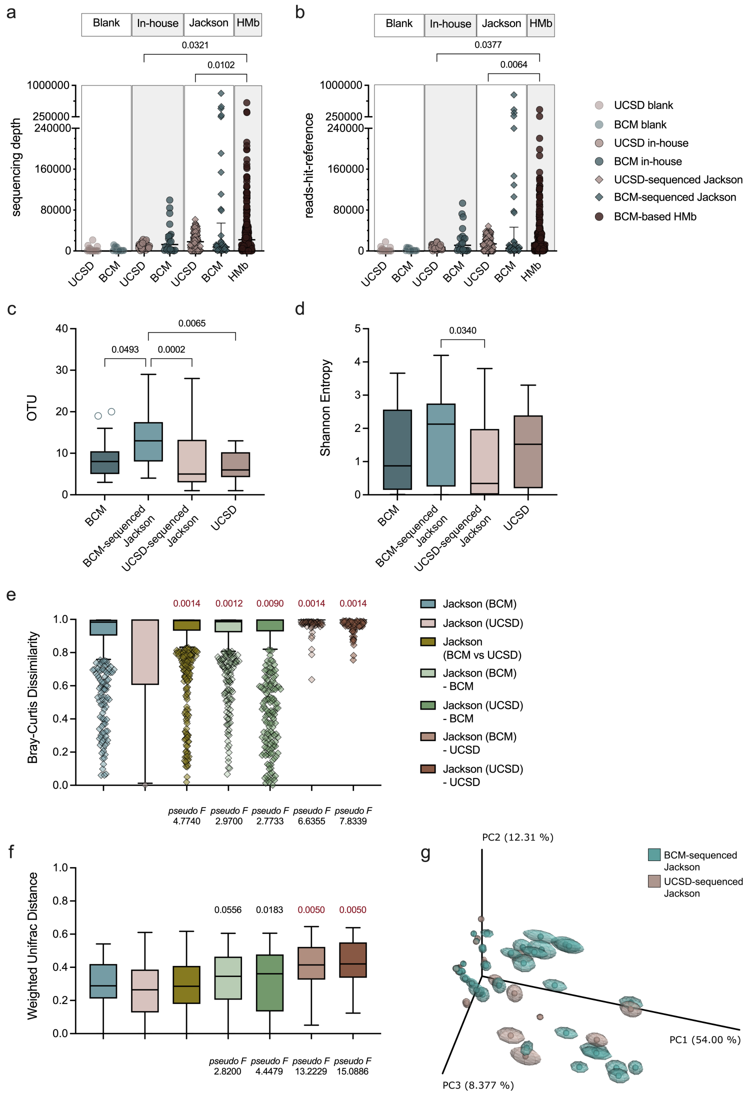
Vaginal microbial dynamics and pathogen colonization in a humanized microbiota mouse model

Marlyd E. Mejia, Vicki Mercado-Evans, Jacob J. Zulk, Samantha Ottinger, Korinna Ruiz, Mallory B. Ballard, Stephanie W. Fowler, Robert A. Britton, and Kathryn A. Patras

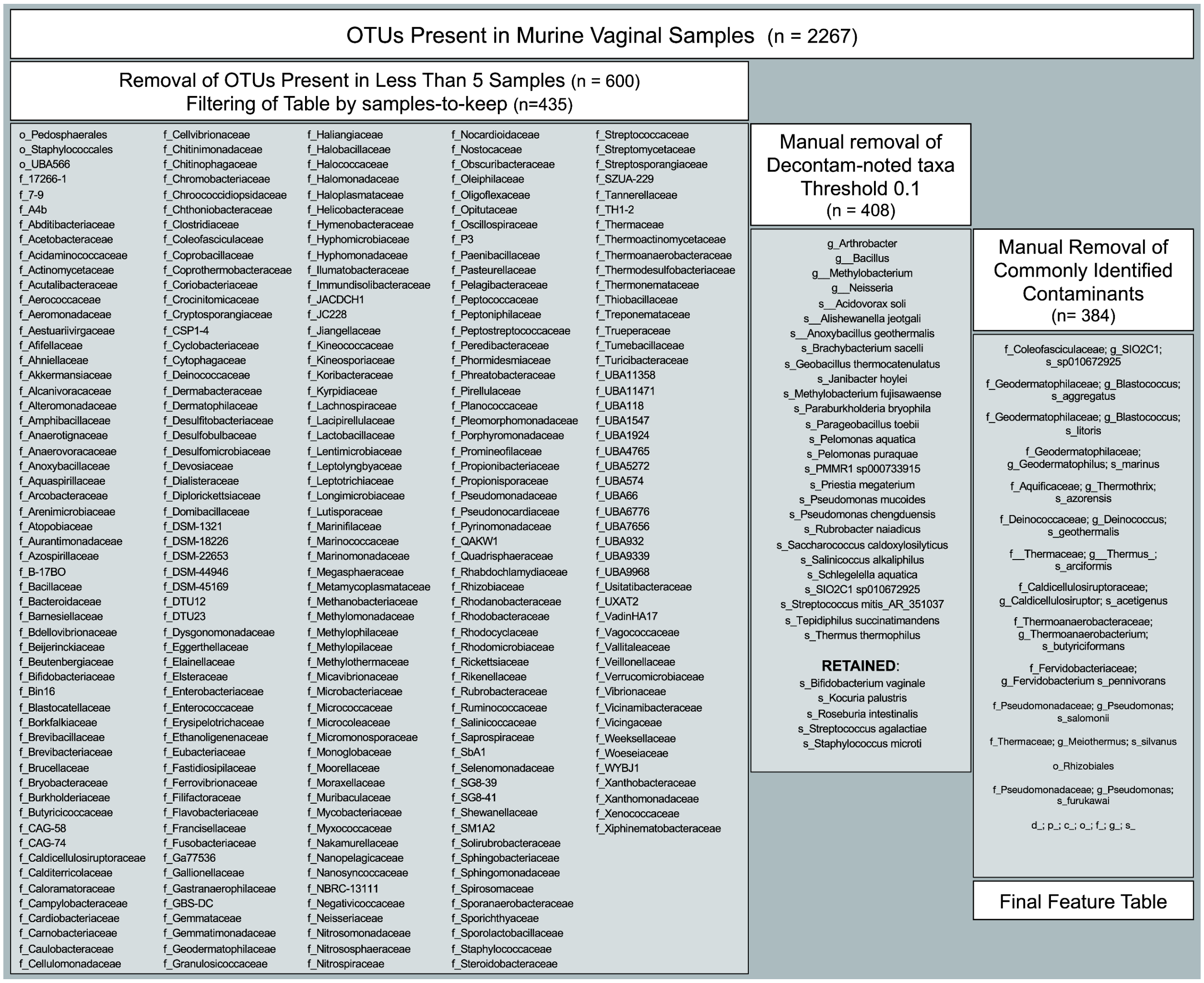
Contents:

Supplementary Figures 1-7

Supplementary Tables 1-2



Supplementary Figure 1. Vaginal microbiome data disaggregated by sequencing site. Sequencing output data from vaginal swabs processed in separate sequencing pipelines were compared. **(a)** Read depth and **(b)** reads matched post deblur to the internal Qiime2 reference database for each mouse colony sequenced at BCM or UCSD. Blanks are included per sequencing site (not included in statistical analysis). **(c)** Observed OTUs and **(d)** Shannon entropy of vaginal swab samples. Colony instability and inter-site variation by **(e)** Bray-Curtis or **(f)** weighted normalized UniFrac distances in comparison to either Jackson mice sequenced at BCM or UCSD. **(g)** PCoA of vaginal communities clustered by weighted normalized UniFrac distances between Jackson colonies. Centroids were determined by jackknifed rarefaction (100 reads). Each symbol (a-b, g) represents a unique vaginal swab community/sample with median and interquartile ranges (a-b). Tukey's boxplots are displayed (c-f). Data were analyzed by Kruskal-Wallis with a Dunn's multiple comparison test excluding blanks (a-d) or PERMANOVA followed by PERMDISP (e-f). PERMANOVA *P* values ($P < 0.1$) are colored in red if PERMDISP was also statistically significant (<0.05). All other statistically significant *P* values are reported. Adjusted *P* values < 0.1 are reported for e-f.



OTUs Present in Murine Vaginal Samples (n = 2267)

Removal of OTUs Present in Less Than 5 Samples (n = 600)
Filtering of Table by samples-to-keep (n=435)

o_Pedosphaerales
o_Staphylococcales
o_UBA566
f_17266-1
f_7-9
f_A4b
f_Abditibacteriaceae
f_Acetobacteraceae
f_Acidaminococcaceae
f_Actinomycetaceae
f_Acutalibacteraceae
f_Aerococcaceae
f_Aeromonadaceae
f_Aestuariivirgaceae
f_Afifellaceae
f_Ahniellaceae
f_Akkermansiaceae
f_Alcanivoracaceae
f_Alteromonadaceae
f_Amphibacillaceae
f_Anaerotignaceae
f_Anaerovoracaceae
f_Anoxybacillaceae
f_Aquaspirillaceae
f_Arcobacteraceae
f_Arenimicrobiaceae
f_Atopobiaceae
f_Aurantimonadaceae
f_Azospirillaceae
f_B-17BO
f_Bacillaceae
f_Bacteroidaceae
f_Barnesiellaceae
f_Bdellovibrionaceae
f_Beijerinckiaceae
f_Beutenbergiaceae
f_Bifidobacteriaceae
f_Bin16
f_Blastocatellaceae
f_Borkfalkiaceae
f_Brevibacillaceae
f_Brevibacteriaceae
f_Brucellaceae
f_Bryobacteraceae
f_Burkholderiaceae
f_Butyricicoccaceae
f_CAG-58
f_CAG-74
f_Caldicellulosiruptoraceae
f_Calditerricolaceae
f_Caloramatoraceae
f_Campylobacteraceae
f_Cardiobacteriaceae
f_Carnobacteriaceae
f_Caulobacteraceae
f_Cellulomonadaceae
f_Cellvibrionaceae
f_Chitinimonadaceae
f_Chitinophagaceae
f_Chromobacteriaceae
f_Chroococcidiopsidaceae
f_Chthoniobacteraceae
f_Clostridiaceae
f_Coleofasciculaceae
f_Coprobacillaceae
f_Coprothermobacteraceae
f_Coriobacteriaceae
f_Crocinitomicaceae
f_Cryptosporangiaceae
f_CSP1-4
f_Cyclobacteriaceae
f_Cytophagaceae
f_Deinococcaceae
f_Dermabacteraceae
f_Dermatophilaceae
f_Desulfitobacteriaceae
f_Desulfobulbaceae
f_Desulfomicrobiaceae
f_Devosiaceae
f_Dialisteraceae
f_Diplorickettsiaceae
f_Domibacillaceae
f_DSM-1321
f_DSM-18226
f_DSM-22653
f_DSM-44946
f_DSM-45169
f_DTU12
f_DTU23
f_Dysgonomonadaceae
f_Eggerthellaceae
f_Elainellaceae
f_Elsteraceae
f_Enterobacteriaceae
f_Enterococcaceae
f_Erysipelotrichaceae
f_Ethanoligenenaceae
f_Eubacteriaceae
f_Fastidiosipilaceae
f_Ferrovibrionaceae
f_Filifactoraceae
f_Flavobacteriaceae
f_Francisellaceae
f_Fusobacteriaceae
f_Ga77536
f_Gallionellaceae
f_Gastranaerophilaceae
f_GBS-DC
f_Gemmataceae
f_Gemmatimonadaceae
f_Geodermatophilaceae
f_Granulosicoccaceae
f_Haliangiaceae
f_Halobacillaceae
f_Halococcaceae
f_Halomonadaceae
f_Haloplasmataceae
f_Helicobacteraceae
f_Hymenobacteraceae
f_Hyphomicrobiaceae
f_Hyphomonadaceae
f_Ilumatobacteraceae
f_Immundisolibacteraceae
f_JACDCH1
f_JC228
f_Jiangellaceae
f_Kineococcaceae
f_Kineosporiaceae
f_Koribacteraceae
f_Kyrpidiaceae
f_Lachnospiraceae
f_Lacipirellulaceae
f_Lactobacillaceae
f_Lentimicrobiaceae
f_Leptolyngbyaceae
f_Leptotrichiaceae
f_Longimicrobiaceae
f_Lutisporaceae
f_Marinifilaceae
f_Marinococcaceae
f_Marinomonadaceae
f_Megasphaeraceae
f_Metamycoplasmataceae
f_Methanobacteriaceae
f_Methylomonadaceae
f_Methylophilaceae
f_Methylopilaceae
f_Methylothermaceae
f_Micavibrionaceae
f_Microbacteriaceae
f_Micrococcaceae
f_Microcoleaceae
f_Micromonosporaceae
f_Monoglobaceae
f_Moorellaceae
f_Moraxellaceae
f_Muribaculaceae
f_Mycobacteriaceae
f_Myxococcaceae
f_Nakamurellaceae
f_Nanopelagicaceae
f_Nanosyncoccaceae
f_NBRC-13111
f_Negativicoccaceae
f_Neisseriaceae
f_Nitrosomonadaceae
f_Nitrososphaeraceae
f_Nitrospiraceae
f_Nocardioidaceae
f_Nostocaceae
f_Obscuribacteraceae
f_Oleiphilaceae
f_Oligoflexaceae
f_Opitutaceae
f_Oscillospiraceae
f_P3
f_Paenibacillaceae
f_Pasteurellaceae
f_Pelagibacteraceae
f_Peptococcaceae
f_Peptoniphilaceae
f_Peptostreptococcaceae
f_Peredibacteraceae
f_Phormidesmiaceae
f_Phreatobacteraceae
f_Pirellulaceae
f_Planococcaceae
f_Pleomorphomonadaceae
f_Porphyromonadaceae
f_Promineofilaceae
f_Propionibacteriaceae
f_Propionisporaceae
f_Pseudomonadaceae
f_Pseudonocardiaceae
f_Pyrinomonadaceae
f_QAKW1
f_Quadrisphaeraceae
f_Rhabdochlamydiaceae
f_Rhizobiaceae
f_Rhodanobacteraceae
f_Rhodobacteraceae
f_Rhodocyclaceae
f_Rhodomicrobiaceae
f_Rickettsiaceae
f_Rikenellaceae
f_Rubrobacteraceae
f_Ruminococcaceae
f_Salinicoccaceae
f_Saprospiraceae
f_SbA1
f_Selenomonadaceae
f_SG8-39
f_SG8-41
f_Shewanellaceae
f_SM1A2
f_Solirubrobacteraceae
f_Sphingobacteriaceae
f_Sphingomonadaceae
f_Spirosomaceae
f_Sporanaerobacteraceae
f_Sporichthyaceae
f_Sporolactobacillaceae
f_Staphylococcaceae
f_Steroidobacteraceae
f_Streptococcaceae
f_Streptomycetaceae
f_Streptosporangiaceae
f_SZUA-229
f_Tannerellaceae
f_TH1-2
f_Thermaceae
f_Thermoactinomycetaceae
f_Thermoanaerobacteraceae
f_Thermodesulfobacteriaceae
f_Thermonemataceae
f_Thiobacillaceae
f_Treponemataceae
f_Trueperaceae
f_Tumebacillaceae
f_Turicibacteraceae
f_UBA11358
f_UBA11471
f_UBA118
f_UBA1547
f_UBA1924
f_UBA4765
f_UBA5272
f_UBA574
f_UBA66
f_UBA6776
f_UBA7656
f_UBA932
f_UBA9339
f_UBA9968
f_Usitatibacteraceae
f_UXAT2
f_VadinHA17
f_Vagococcaceae
f_Vallitaleaceae
f_Veillonellaceae
f_Verrucomicrobiaceae
f_Vibrionaceae
f_Vicinamibacteraceae
f_Vicingaceae
f_Weeksellaceae
f_Woeseiaceae
f_WYBJ1
f_Xanthobacteraceae
f_Xanthomonadaceae
f_Xenococcaceae
f_Xiphinematobacteraceae

Manual removal of
Decontam-noted taxa
Threshold 0.1
(n = 408)

g_Arthrobacter
g__Bacillus
g__Methylobacterium
g__Neisseria
s__Acidovorax soli
s__Alishewanella jeotgali
s__Anoxybacillus geothermalis
s_Brachybacterium sacelli
s_Geobacillus thermocatenulatus
s_Janibacter hoylei
s_Methylobacterium fujisawaense
s_Paraburkholderia bryophila
s_Parageobacillus toebii
s_Pelomonas aquatica
s_Pelomonas puraquae
s_PMMR1 sp000733915
s_Priestia megaterium
s_Pseudomonas mucoides
s_Pseudomonas chengduensis
s_Rubrobacter naiadicus
s_Saccharococcus caldoxylosilyticus
s_Salinicoccus alkaliphilus
s_Schlegelella aquatica
s_SIO2C1 sp010672925
s_Streptococcus mitis_AR_351037
s_Tepidiphilus succinatimandens
s_Thermus thermophilus

RETAINED:
s_Bifidobacterium vaginale
s_Kocuria palustris
s_Roseburia intestinalis
s_Streptococcus agalactiae
s_Staphylococcus microti

Manual Removal of
Commonly Identified
Contaminants
(n= 384)

f_Coleofasciculaceae; g_SIO2C1; s_sp010672925

f_Geodermatophilaceae; g_Blastococcus; s_aggregatus

f_Geodermatophilaceae; g_Blastococcus; s_litoris

f_Geodermatophilaceae; g_Geodermatophilus; s_marinus

f_Aquificaceae; g_Thermothrix; s_azorensis

f_Deinococcaceae; g_Deinococcus; s_geothermalis

f__Thermaceae; g__Thermus_; s_arciformis

f_Caldicellulosiruptoraceae; g_Caldicellulosiruptor; s_acetigenus

f_Thermoanaerobacteraceae; g_Thermoanaerobacterium; s_butyriciformans

f_Fervidobacteriaceae; g_Fervidobacterium s_pennivorans

f_Pseudomonadaceae; g_Pseudomonas; s_salomonii

f_Thermaceae; g_Meiothermus; s_silvanus

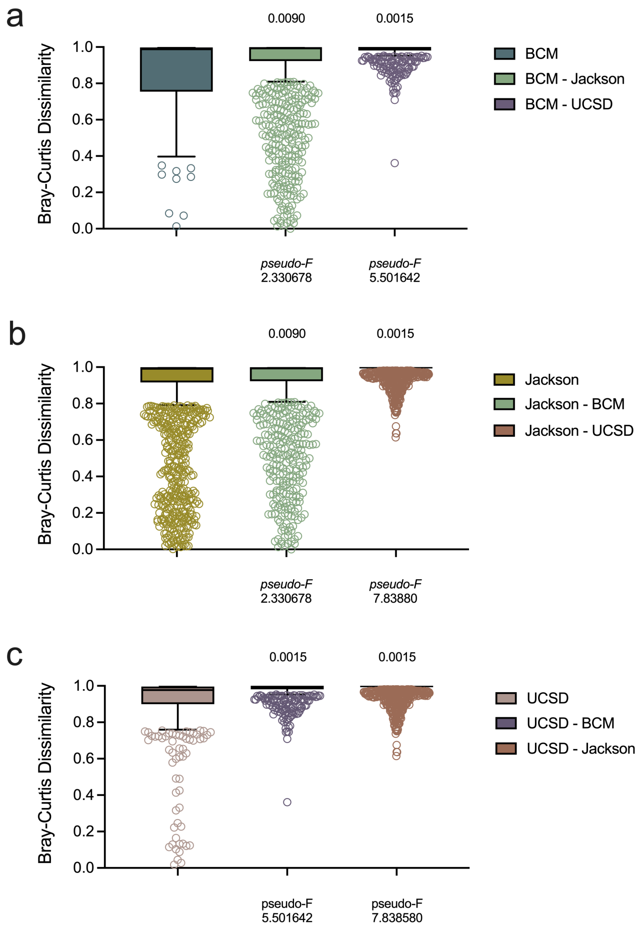
o_Rhizobiales

f_Pseudomonadaceae; g_Pseudomonas; s_furukawai

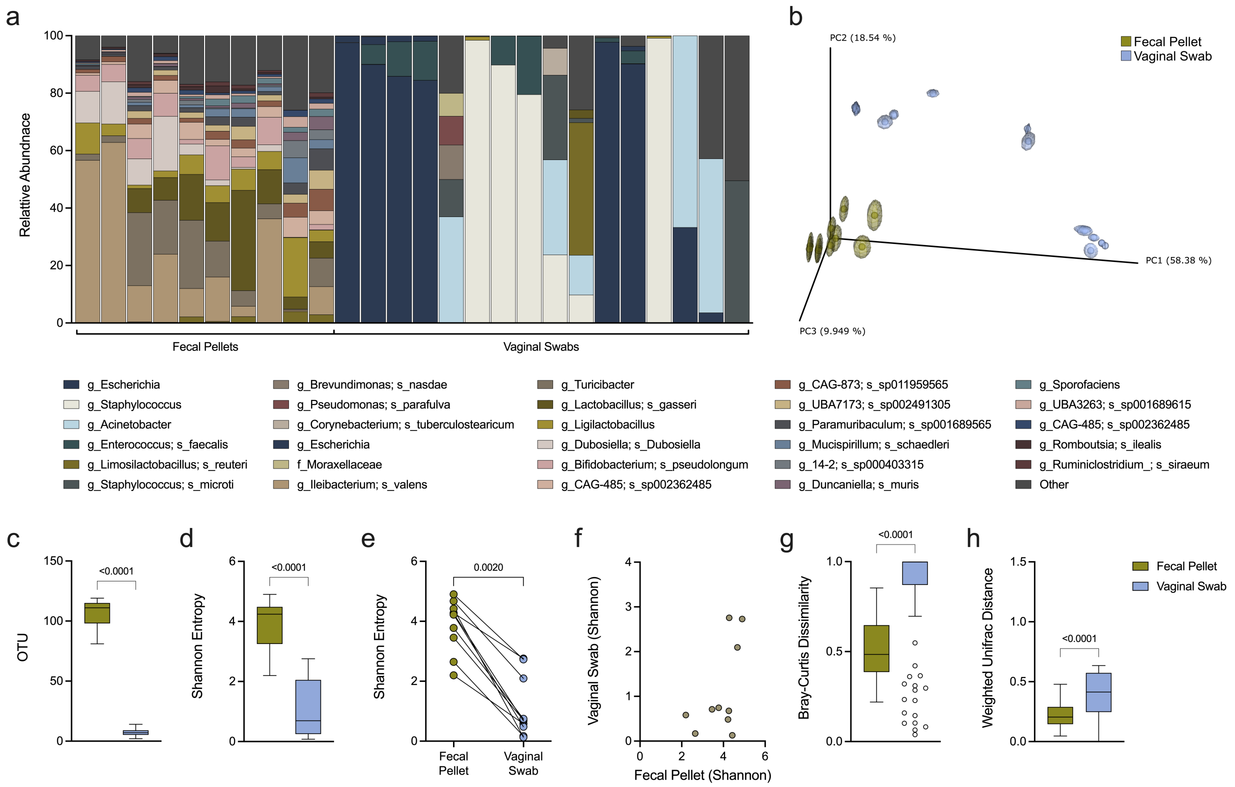
d_; p_; c_; o_; f_; g_; s_

Final Feature Table

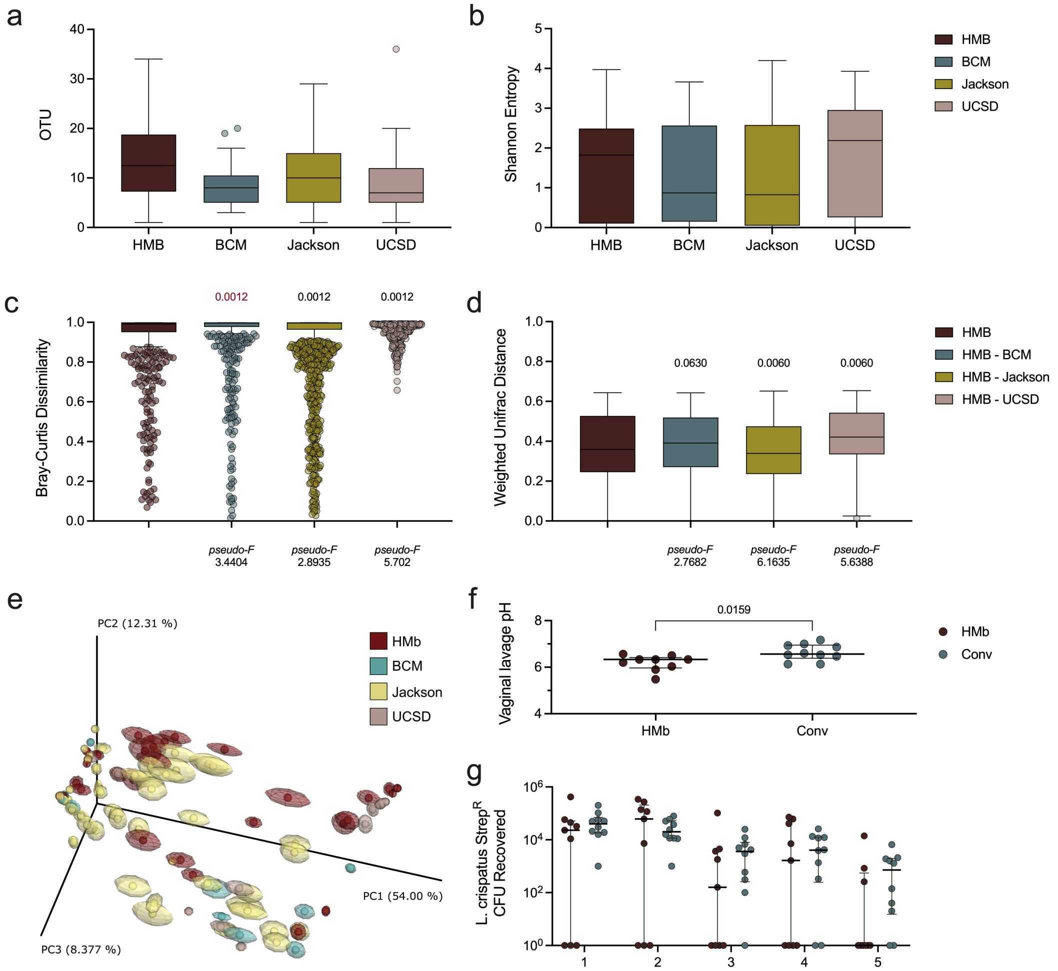
Supplementary Figure 2. Flow chart of contaminant sequence removal to generate feature tables. All studies were merged prior to taxonomic assignment. In the resulting feature table, OTUs that appeared in less than 5 samples were removed in Qiime2 followed by removal of contaminants using Decontam in R. The table was re-imported into Qiime2 for the manual removal of known contaminants.



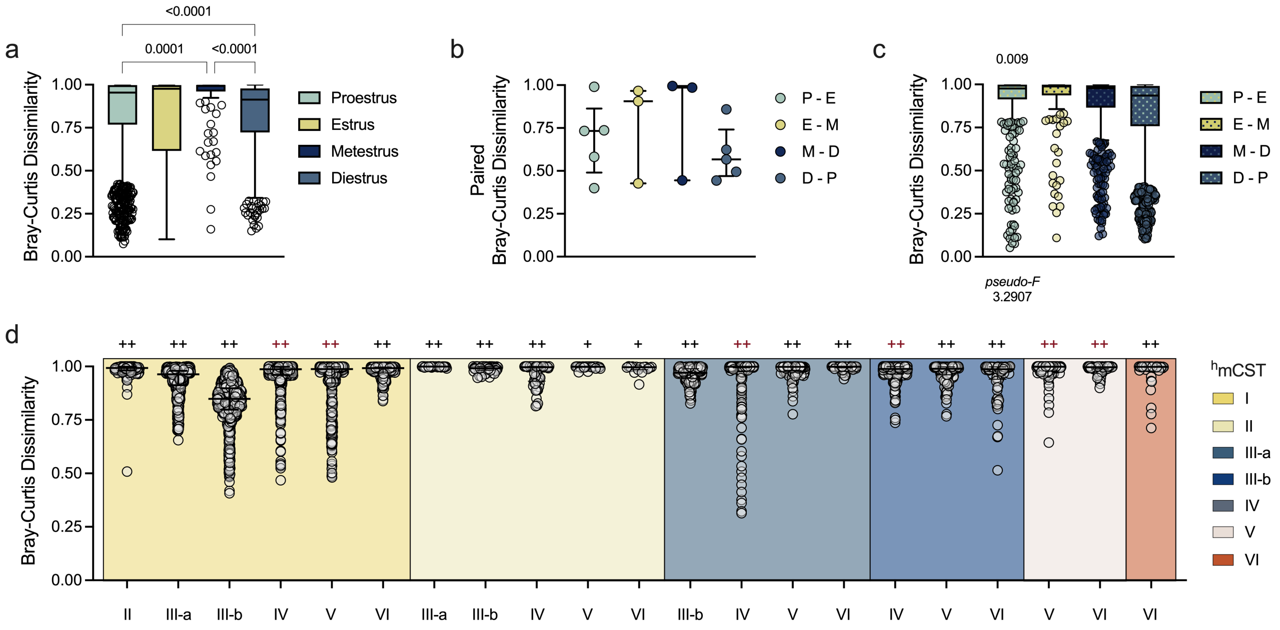
Supplementary Figure 3. The microbial composition of the murine vaginal tract varies between vivaria. Vaginal swabs from mouse colonies raised at BCM (*n* = 21), Jackson Lab (*n* = 30), and UCSD (*n* = 12, mice repeatedly swabbed) were collected and sequenced over the 16S rRNA V4 region. Colony instability and inter-site variation (Bray-Curtis dissimilarity) in comparison to **(a)** BCM, **(b)** Jackson Labs, and **(c)** UCSD. Tukey's boxplots are displayed (a-c). Data were analyzed by PERMANOVA followed by PERMDISP (a-c). PERMANOVA *P* values (*P* < 0.1) are colored in red if PERMDISP was also statistically significant (*P* < 0.05). Adjusted *P* values < 0.1 are reported.



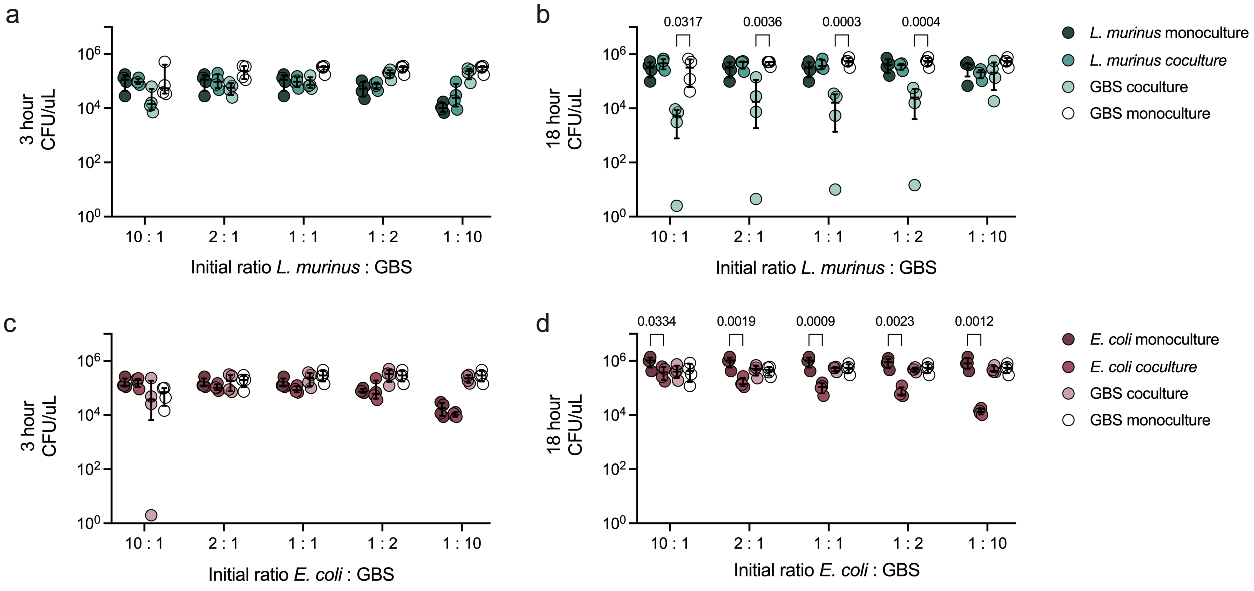
Supplementary Figure 4. $^{\text{HMb}}$mice have tissue-specific microbial compositions. Vaginal swab and fecal pellet sets were collected from a cohort of $^{\text{HMb}}$mice (*n* = 16) **(a)** Microbial compositions of fecal pellets (left) and vaginal swabs (right) where the first ten vaginal communities are matched to fecal communities in corresponding order. **(b)** PCoA of vaginal and fecal communities clustered by weighted normalized UniFrac distances. Centroids were determined by jackknifed rarefaction (100 reads). **(c)** Observed OTUs, **(d)** Shannon entropy, and **(e)** pairwise Shannon entropy (*n* = 10) between vaginal and fecal communities. **(f)** Correlation between Shannon entropy in vaginal and fecal samples within mice (*n* = 10). **(g)** Bray-Curtis or **(h)** weighted UniFrac distances from both sample types. Tukey's boxplots (c-d, f-g) or individual samples or comparisons marked as columns (a) or symbols (b, e) are displayed. Data were statistically analyzed by Mann-Whitney (c-d, g-h) or Wilcoxon Rank Sum test (e), or Spearman Correlation (f). Statistically significant *P* values are reported.



Supplementary Figure 5. ^{HMb}mice display unique and variable vaginal microbiota compared to conventional mice. ^{HMb}mice vaginal compositions were compared to the vaginal microbiota of the conventionally colonized mice from **Fig. 1**. **(a)** Observed OTUs and **(b)** Shannon entropy of vaginal swab samples. HMb colony instability and inter-colony variation by **(c)** Bray-Curtis and **(d)** weighted normalized UniFrac distances in comparison to BCM, Jackson Labs, and UCSD. **(e)** PCoA of vaginal communities clustered by weighted normalized UniFrac distances between colonies. Centroids were determined by jackknifed rarefaction (100 reads). **(f)** Vaginal pH of ^{HMb}mice and conventionally colonized BCM mice determined from vaginal lavage. **(g)** Duration and abundance of Lactobacillus colonization in ^{HMb}mice and conventionally colonized BCM mice. Experiments were performed as one independent replicate (f-g). Tukey's boxplots (a-d) and individual samples marked as symbols (e) or symbols with median and interquartile ranges (f-g) are displayed. Data were statistically analyzed with Kruskal-Wallis with Dunn's multiple comparisons test (a-b), by PERMANOVA followed by PERMDISP (c-d), or Mann-Whitney (f) with corrections for multiple comparisons using the two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli and a false discovery rate (<0.05) (g). PERMANOVA *P* values ($P < 0.1$) are colored in red if PERMDISP was also statistically significant ($P < 0.05$). Statistically significant *P* values are reported. Adjusted *P* values < 0.1 are reported for c-d, g.



Supplementary Figure 6. Vaginal microbiota dynamics across estrous stages and between ^{h}mCST in ^{HMb}mice. Vaginal swabs were collected at baseline and assigned the host's estrous stage at time of collection. Bray-Curtis dissimilarity of microbial compositions **(a)** among samples categorized in the same estrous stage (*n* = 35), **(b)** between paired, consecutive samples collected from individual mice, **(c)** between unpaired samples from sequential estrous stages, or **(d)** between ^{h}mCSTs (*n* = 183 samples). ^{h}mCSTs labeled below are compared to the ^{h}mCST of the corresponding background color. **(d)** Tukey's boxplots (a,c) or individual comparisons marked as symbols with median and interquartile ranges (b,d) are displayed. Data were statistically analyzed by Kruskal-Wallis with Dunn's multiple comparisons test (a-b) or PERMANOVA followed by PERMDISP (c-d). PERMANOVA P values ($P < 0.1$) are colored in red if PERMDISP was also statistically significant ($P < 0.05$). All statistically significant P values are reported. $^{+}P < 0.03$; $^{++}P < 0.003$.



Supplementary Figure 7. Viable CFU recovered from GBS competition experiments with *L. murinus* and *E. coli*. GBS was cocultured with either **(a, b)** *L. murinus* or **(c, d)** *E. coli* at increasing concentrations of GBS. Viable CFU of all microbes in coculture as well as monoculture controls were quantified at **(a, e)** 3 hours and **(b, d)** 18 hours. Individual experimental replicates marked as symbols with median and interquartile ranges are displayed. Comparisons between monoculture and coculture were analyzed by 2-way ANOVA with Šídák's multiple comparison test; statistically significant *P* values are shown.

Supplementary Table 1. *Ligilactobacillus* species candidates for $^{\mathrm{HMb}}$mice-associated OTUs cross-referenced through BLAST[1]

Taxonomy	OTU[2]	16S v4 Region Sequence	BLAST result
g_ Ligilactobacillus; s_	RS-GCF-910574405.1-NZ-CAJTAQ010000093.1	TACGTAGGTGGCAAGCGTTATCCGGATTTATTGGGCGTAAAGGGAACGCAGGCGGTCTTTTAAGTCTGATGTGAAAGCCTTCGGCTTAACCGGAGTAGTGCATTGGAAACTGGGAGACTTGAGTGCAGAAGAGGAGAGTGGAACTCCATG	Ligilactobacillus animalis Ligilactobacillus murinus Uncultured bacterium (100% identity)
	MJ005-barcode27-umi40855bins-ubs-6	TACGTAGGTGGCAAGCGTTATCCGGATTTATTGGGCGTAAAGGGAACGCAGGCGGTCTTTTAAGTCTGATGTGAAAGCCTTCGGCTTAACCGGAGTAGTGCATTGGAAACTGGGAGGCTTGAGTGCAGAAGAGGAGAGTGGAACTCCATG	Uncultured bacterium (100% identity) Uncultured Firmicutes Ligilactobacillus animalis Ligilactobacillus murinus Ligilactobacillus faecis (99.3% identity)
	MJ009-1-barcode35-umi86035bins-ubs-4	TACGTAGGTGGCAAGCGTTATCCGGATTTATTGGGCGTAAAGGGAACGCAGGCGGTCTTTTAATTCTGATGTGAAAGCCTTCGGCTTAACCGGAGTAGTGCATTGGAAACTGGGAGACTTGAGTGCAGAAGAGGAGAGTGGAACTCCATG	Ligilactobacillus animalis Ligilactobacillus murinus Ligilactobacillus faecis Uncultured bacterium (99.3% identity)
	MJ009-2-barcode52-umi43642bins-ubs-14	TACGTAGGTGGCAAGCGTTATCCGGATTTATTGGGCGTAAAGAGAACGCAGGCGGTCTTTTAAGTCTGATGTGAAAGCCTTCGGCTTAACCGGAGTAGTGCATTGGAAACTGGGAGACTTGAGTGCAGAAGAGGAGAGTGGAACTCCATG	Ligilactobacillus murinus Uncultured bacterium (100% identity)
	RS-GCF-004793535.1-NZ-SRYK01000136.1	TACGTAGGTGGCAAGCGTTATCCGGATTTATTGGGCGTAAAGGGAACGCAGGCGGTCTTTTAAGTCTGATGTGAAAGCCTTCGGCTTACCCGGAGTAGTGCATTGGAAACTGGGAGACTTGAGTGCAGAAGAGGAGAGTGGAACTCCATG	Uncultured bacterium (100% identity) Ligilactobacillus animalis Ligilactobacillus murinus (99.3% identity)

[1] Timestamp for BLAST search is Sep. 5, 2023 at 11:40:08.

[2] Top two boldened OTUs are the first and second most abundant OTUs shown in **Fig. 4a**. Other *Ligilactobacillus* OTUs in the study, but not displayed Fig. 4a, are noted below the grey bar.

Supplementary Table 2. Statistics[1] on ^{h}mCST compositional comparisons

			Bray-Curtis						Weighted normalized UniFrac					
			PERMANOVA			PERMDISP			PERMANOVA			PERMDISP		
Group 1	Group 2	Sample size	pseudo-F	p-value	q-value	F-value	p-value	q-value	pseudo-F	p-value	q-value	F-value	p-value	q-value
I	II	50	12.433	0.001	0.001	0.908	0.522	0.693	14.459	0.001	0.001	0.017	0.868	0.868
I	III-a	61	38.618	0.001	0.001	0.310	0.567	0.700	441.093	0.001	0.001	5.428	0.019	0.033
I	III-b	64	28.330	0.001	0.001	4.103	0.051	0.153	187.833	0.001	0.001	6.696	0.013	0.027
I	IV	68	22.759	0.001	0.001	106.351	0.001	0.005	28.772	0.001	0.001	147.736	0.001	0.003
I	V	58	25.598	0.001	0.001	13.192	0.001	0.005	85.269	0.001	0.001	3.993	0.050	0.081
I	VI	52	21.418	0.001	0.001	0.026	0.912	0.959	34.302	0.001	0.001	0.155	0.703	0.738
II	III-a	17	6.906	0.001	0.001	0.107	0.641	0.748	331.763	0.002	0.003	3.018	0.003	0.007
II	III-b	20	6.335	0.001	0.001	0.001	0.968	0.968	42.674	0.003	0.004	0.944	0.386	0.450
II	IV	24	1.866	0.002	0.002	19.327	0.099	0.231	2.799	0.017	0.018	18.220	0.003	0.007
II	V	14	4.439	0.004	0.004	0.347	0.528	0.693	42.290	0.005	0.006	2.401	0.074	0.111
II	VI	8	5.299	0.024	0.024	0.216	0.459	0.689	14.915	0.030	0.030	0.358	0.338	0.418
III-a	III-b	31	18.897	0.001	0.001	0.436	0.349	0.590	40.802	0.001	0.001	20.529	0.001	0.003
III-a	IV	35	7.606	0.001	0.001	20.934	0.001	0.005	23.878	0.001	0.001	114.700	0.001	0.003
III-a	V	25	13.372	0.001	0.001	1.666	0.091	0.231	1213.320	0.001	0.001	0.271	0.443	0.490
III-a	VI	19	11.841	0.001	0.001	0.010	0.913	0.959	529.781	0.001	0.001	1.181	0.247	0.341
III-b	IV	38	8.454	0.001	0.001	29.356	0.001	0.005	12.470	0.001	0.001	59.232	0.001	0.003
III-b	V	28	13.076	0.001	0.001	0.944	0.272	0.519	174.170	0.001	0.001	11.502	0.001	0.003
III-b	VI	22	9.870	0.002	0.002	0.671	0.365	0.590	73.378	0.001	0.001	4.886	0.019	0.033
IV	V	32	5.409	0.001	0.001	16.328	0.002	0.008	11.582	0.001	0.001	90.609	0.001	0.003
IV	VI	26	4.015	0.001	0.001	35.181	0.007	0.025	4.942	0.003	0.004	33.970	0.001	0.003
V	VI	16	7.454	0.001	0.001	2.036	0.138	0.290	61.996	0.001	0.001	1.054	0.260	0.341

[1] Pair-wise PERMANOVA and PERMDISP were performed with 999 permutations using Qiime2 package “qiime diversity beta-group-significance”.