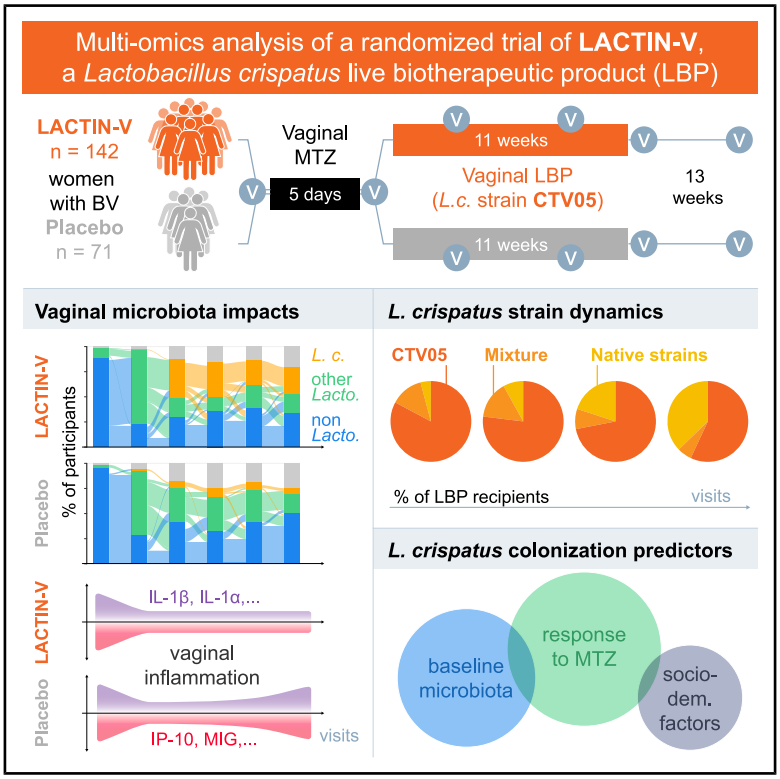
Clinical and Translational Report

Cell Host & Microbe

Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial

Graphical abstract



Highlights

- LACTIN-V recipients achieve higher rates of *Lactobacillus crispatus* dominance
- The LBP strain predominates, but native strains increasingly replace it over time
- LBP efficacy differs in women with different pre-treatment vaginal microbiota
- *L. crispatus* dominance in the LBP arm varies by biological and sociodemographic factors

Authors

Seth M. Bloom, Laura Symul, Joseph Elsherbini, ..., Craig R. Cohen, Susan P. Holmes, Douglas S. Kwon

Correspondence

dkwon@mgh.harvard.edu

In brief

Bloom, Symul, et al. show that a vaginal live biotherapeutic product (LBP) containing the health-associated bacterium *Lactobacillus crispatus* prevents bacterial vaginosis recurrence and reduces vaginal inflammation. Colonization is usually driven by the administered strain, but endogenous strains sometimes replace it. Success varies depending on pre-treatment vaginal microbiota composition and host factors.

Bloom et al., 2026, Cell Host & Microbe *34*, 1–17
April 8, 2026 © 2026 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.
https://doi.org/10.1016/j.chom.2026.03.003

CellPress

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

Cell Host & Microbe

CellPress

Clinical and Translational Report

Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial

Seth M. Bloom,[1,2,3,10] Laura Symul,[4,10] Joseph Elsherbini,[1] Jiawu Xu,[1] Salina Hussain,[1] Johnathan Shih,[1] Ashley Sango,[1] Caroline M. Mitchell,[1,3,5,6] Anke Hemmerling,[7] Thomas P. Parks,[8] Aditi Kannan,[1] Fatima A. Hussain,[1,3,11] Craig R. Cohen,[7] Susan P. Holmes,[9] and Douglas S. Kwon[1,2,3,12,*]

[1]Ragon Institute of Mass General Brigham, MIT, and Harvard, Cambridge, MA 02139, USA
[2]Infectious Diseases Division, Massachusetts General Hospital, Boston, MA 02114, USA
[3]Harvard Medical School, Boston, MA 02115, USA
[4]Institute of Statistics, Biostatistics and Actuarial Sciences, Louvain Institute of Data Analysis and Modelling, UCLouvain, 1348 Louvain-la-Neuve, Belgium
[5]Department of Obstetrics & Gynecology, Massachusetts General Hospital, Boston, MA 02114, USA
[6]Vincent Center for Reproductive Biology, Massachusetts General Hospital, Boston, MA 02114, USA
[7]Department of Obstetrics, Gynecology, & Reproductive Sciences, University of California, San Francisco, San Francisco, CA 94158, USA
[8]Osel, Inc., Mountain View, CA 94043, USA
[9]Department of Statistics, Stanford University, Stanford, CA 94305, USA
[10]These authors contributed equally
[11]Present address: Department of Biology, Tufts University, Medford, MA 02155, USA
[12]Lead contact
*Correspondence: dkwon@mgh.harvard.edu
https://doi.org/10.1016/j.chom.2026.03.003

SUMMARY

Bacterial vaginosis (BV) affects >25% of women worldwide and often recurs after standard-of-care metronidazole (MTZ) treatment. LACTIN-V, a live biotherapeutic product (LBP) containing *Lactobacillus crispatus* strain CTV-05, significantly reduced recurrent BV in a phase 2b clinical trial, but efficacy was incomplete. Here, we characterize microbiota and immune effects using multi-omics and define correlates of treatment success. By week 12, an *L. crispatus*-dominant microbiota was achieved in 30% of LBP recipients compared with 9% of placebo recipients (benefit ratio: 3.31; $p < 0.005$). This is primarily due to CTV-05, but native *L. crispatus* strains are also present and increase over time. Inflammatory cytokines decrease in both arms after MTZ but return to baseline in placebo recipients. Successful *L. crispatus* colonization is associated with pre-MTZ microbiota, baseline inflammatory profiles, post-MTZ bacterial load, and clinical and behavioral variables. These findings elucidate LBP microbiota effects and identify predictors of treatment success, informing improved intervention strategies to advance women's health.

INTRODUCTION

Bacterial vaginosis (BV), a syndrome characterized by *Lactobacillus*-deficient vaginal microbiota, affects 23%–29% of reproductive-age women globally, with an estimated annual economic burden of $4.8 billion.[1,2] BV symptoms include discharge, malodor, pain, and itching, significantly impairing quality of life, self-esteem, and sexual well-being.[3,4] BV is also linked to mucosal inflammation and a higher risk of adverse outcomes, including HIV acquisition, sexually transmitted infections, preterm birth, human papillomavirus infection, and cervical cancer.[5–10] Although BV occurs globally, it disproportionately affects those with lower socioeconomic status and members of racial or ethnic minority groups across diverse settings.[1,11–13] Effectively treating BV is therefore a key objective for improving women's health.[14]

Current first-line BV therapy involves oral or intravaginal antibiotics such as metronidazole (MTZ), which target species in the diverse anaerobic bacterial communities characteristic of BV.[2,14] In most cases, MTZ reduces the abundance of BV-associated anaerobes, leading to the emergence of bacterial communities dominated by *Lactobacillus* species (which are intrinsically MTZ-resistant),[15] but BV frequently recurs post-treatment at rates ranging from >50% within 1 year to >75% within 16 weeks.[16,17] It is hypothesized that MTZ's incomplete efficacy results from causes including failure to eradicate BV-associated bacterial communities or post-MTZ replacement by *Lactobacillus* species prone to reverting to BV.[15,18] Specifically, *Lactobacillus iners* is associated with an increased risk of transition to BV-like states[19–21] and higher rates of adverse health outcomes.[6,22–25] MTZ frequently results in the

All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

establishment of vaginal microbiota dominated by *L. iners* rather than *L. crispatus*,[26–31] providing a strong rationale that therapies promoting *L. crispatus* over *L. iners* may improve outcomes and enhance vaginal health.[15,32,33]

New strategies to promote *L. crispatus* colonization during BV treatment are currently in development, including vaginal microbiome transplants, adjunctive therapies that selectively inhibit *L. iners* or promote *L. crispatus* growth, and *L. crispatus*-containing live biotherapeutic products (LBPs).[32–39] LACTIN-V, a vaginal formulation containing *L. crispatus* strain CTV-05, is the only LBP evaluated in large-scale clinical trials to date.[37,40] Cohen and colleagues reported clinical results of a phase 2b randomized, placebo-controlled, double-blinded study showing that participants who received 11 weeks of LACTIN-V after intravaginal MTZ developed recurrent BV (rBV) at significantly lower rates than placebo.[37] However, rBV remained common, occurring in 30% and 39% of LACTIN-V recipients by weeks 12 and 24, respectively (versus 45% and 54%, respectively, in placebo recipients). Notably, the study characterized rBV clinically, but comprehensive molecular analysis of microbiota and factors associated with *L. crispatus* colonization was not performed.

Here, we analyze samples from that trial to characterize LACTIN-V's effects on vaginal microbiota composition and identify correlates of successful *L. crispatus* colonization by assessing vaginal microbial communities, bacterial strain dynamics, cytokine trends, and demographic/behavioral parameters.[37] Reduction in rBV among LBP recipients corresponded to >3-fold higher rates of achieving *L. crispatus*-dominant microbiota. However, only 30% achieved *L. crispatus* dominance at week 12, consistent with observed incomplete clinical efficacy. The CTV-05 strain accounted for the majority of *L. crispatus* colonization among LBP recipients, but native strains increased over time and sometimes displaced CTV-05. We describe immune correlates of LBP treatment and show that pre-MTZ microbiota, baseline inflammatory profiles, post-MTZ total bacterial load, and selected clinical/behavioral variables were associated with *L. crispatus* dominance in LBP recipients. Our findings reveal how an *L. crispatus* LBP alters the vaginal microbiota to improve BV treatment efficacy and identify factors linked to LBP success that can guide the development of therapies to improve women's health.

RESULTS

Study design and participant characteristics

Women aged 18–45 years with BV completed 5 days of intravaginal MTZ and then were randomized 2:1 to intravaginal LACTIN-V (Osel) or placebo, as previously described (see also STAR Methods).[37] Enrollment occurred at four US sites, and ~50% of participants reported ≥5 prior lifetime BV episodes.[37] LACTIN-V is a powder formulation containing live *L. crispatus* strain CTV-05,[41–43] administered as 2×10^9 colony-forming units per dose. BV was diagnosed at the baseline "pre-MTZ" visit based on at least three of four Amsel criteria and a Nugent score ≥ 4.[44,45] Participants received intravaginal MTZ for 5 days and then were randomized to LBP or placebo within 48 h of antibiotic completion at a "post-MTZ" visit. The first LBP (or placebo) dose was clinician administered at the post-MTZ visit, then self-administered daily for 4 consecutive days, and then twice weekly for 10 weeks (Figure 1A). Vaginal swabs were collected at the pre-MTZ visit, post-MTZ (before product administration), and at weeks 4, 8, 12, and 24 post-randomization. Demographic, clinical, and behavioral data showed no notable between-arm differences.[37] 228 participants (152 LBP and 76 placebo) were enrolled. Samples from 213 (142 LBP and 71 placebo) were available for our analysis, representing 1,156 unique visits.

LBP treatment significantly increased rates of *L. crispatus*-dominant vaginal microbiota

Microbiota composition was determined by bacterial 16S ribosomal RNA gene sequencing. Four samples were excluded due to technical failures. The remaining samples (Table S1; Figure S1A) had a median of 31,177 (IQR 23,774–41,251) analyzable reads. Microbiota composition did not differ between arms at pre-MTZ or post-MTZ visits as assessed by permutational multivariate analysis of variance (PERMANOVA $p = 0.63$ and $p = 0.14$, respectively; Figures S1B and S1C).

We pre-specified two microbiota outcome parameters: ≥50% *L. crispatus* relative abundance ("*L. crispatus* dominant"; a threshold useful for distinguishing outcomes in other studies[6,21]) and ≥50% summed relative abundance of all *Lactobacillus* species ("*Lactobacillus* dominant," including *L. crispatus*). The primary endpoint for our analysis was *L. crispatus* dominance at week 12, the visit corresponding to the trial's primary clinical endpoint. Secondary endpoints included *L. crispatus* dominance at week 24 and total *Lactobacillus* dominance at weeks 12 or 24. Almost all participants had <50% *Lactobacillus* abundance pre-MTZ, consistent with BV diagnosis (Figure 1B; Data S1). Majorities in both arms transitioned to *Lactobacillus* dominance at the post-MTZ visit, largely driven by *L. iners* (52% of participants *L. iners* dominant, 8% *L. jensenii/mulieris* dominant, and 13% *Lactobacillus* dominant with multiple species). Only two participants (1%, one per arm) had post-MTZ *L. crispatus* dominance (Figure 1B). Microbiota composition diverged post-randomization, with 30% of LBP recipients ($n = 37$) and only 9% of placebo recipients ($n = 5$) having *L. crispatus* dominance at week 12, a benefit ratio of 3.4 (95% confidence interval [CI]: 1.4–8.1; $p < 0.005$; Table 1; Figures 1B and 1C). At week 24, 35% of LBP recipients ($n = 39$) versus 8% of placebo recipients ($n = 4$) had *L. crispatus* dominance (benefit ratio 4.5; 95% CI: 1.7–11.9). *L. crispatus* relative abundance at weeks 12 and 24 was bimodal and well-separated by the pre-specified 50% threshold, with <4% (week 12) and <7% (week 24) of participants exhibiting relative abundances between 33% and 67%, indicating that results were robust to alternative thresholds (Figure 1D). LBP did not increase total *Lactobacillus* dominance at week 12 (56% and 48% of LBP and placebo recipients, respectively; benefit ratio: 1.16; 95% CI: 0.85–1.59) but did at week 24 (58% and 33% for LBP and placebo, respectively; benefit ratio: 1.76; 95% CI: 1.15–2.68; Table 1; Figures 1B and 1C). To confirm replication of known relationships between *Lactobacillus* and BV absence,[46] we compared microbiota category to concurrent BV status at each post-randomization visit ($n = 729$ visits; Table S2; Figure 1C). rBV was present at no visits with ≥50% *L. crispatus* colonization

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

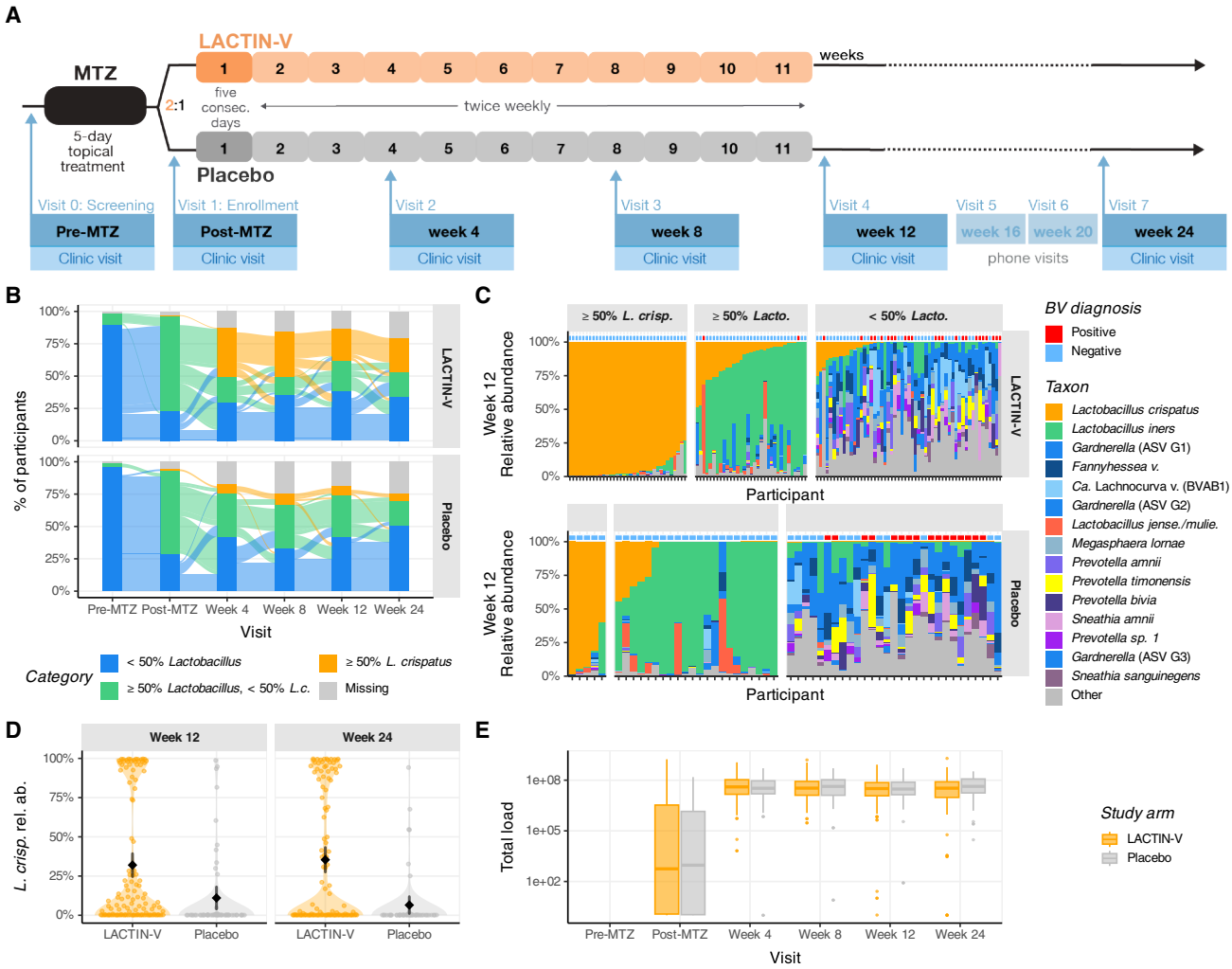


Figure 1. LBP treatment increased the proportion of women with *L. crispatus*-dominant vaginal microbiota

(A) LACTIN-V trial design. See also Table S1 and Figure S1.

(B) Sankey diagram showing transitions between microbiota categories: *L. crispatus* dominance (≥50% *L. crispatus* relative abundance), *Lactobacillus* (non-*crispatus*) dominance (≥50% *Lactobacillus*, <50% *L. crispatus*), or non-*Lactobacillus* dominance (<50% *Lactobacillus*). See also Data S1.

(C) Week 12 microbiota composition for LBP and placebo recipients, grouped by (B) categories, with BV diagnosis at the top. See also Table S2.

(D) *L. crispatus* relative abundances by arm at weeks 12 and 24. Black dots and whiskers show the mean and 95% CI.

(E) Total bacterial load by quantitative PCR (qPCR) at each scheduled visit. Boxplots (here and in subsequent figures) show the median (middle horizontal line), 25th and 75th percentiles (lower and upper box boundaries, interquartile range or "IQR"), values within 1.5 times the IQR (whiskers), and individual values outside that range (dots).

(0 of 209) and only 2 of 194 visits with <50% *L. crispatus* but ≥50% total *Lactobacillus*. rBV was present in over half (164 of 326) of non-*Lactobacillus*-dominant visits, confirming that *Lactobacillus* dominance in our cohort was specific for BV absence, while non-*Lactobacillus* dominance was sensitive but less specific for BV presence.

Total vaginal bacterial load was assessed via quantitative PCR (qPCR) at post-MTZ and subsequent visits. Median load was significantly lower and had a wider range post-MTZ (median $10^{2.69}$ copies/swab; IQR $10^{0.08}$–$10^{6.44}$) than at subsequent visits (median $10^{7.52}$ copies/swab; IQR $10^{7.09}$–$10^{7.93}$), consistent with antibiotic-mediated microbiota biomass depletion followed by repopulation (Figure 1E). Post-MTZ microbiota composition and bacterial load did not significantly correlate, as assessed using the RV coefficient, a multi-table multivariate generalization of the squared Pearson correlation[47,48] (RV coefficient = 0.02, p value > 0.1).

Microbiota composition established early during treatment tended to persist

To characterize microbiota trajectories with greater resolution and reduce data dimensionality for subsequent analyses, we summarized microbiota composition using "topic" mixtures.[49,50] Compared with clustering or categorizing samples into community state types (CSTs),[51] cervicotypes,[6] or sub-CSTs,[52] this method provides probabilistic weightings, enabling more accurate descriptions of taxonomic composition and longitudinal dynamics.[50] In our cohort, reference CST assignment performed

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

Table 1. LBP treatment significantly increased rates of *L. crispatus*-dominant vaginal microbiota

Taxon	Week	Microbiota endpoint	LACTIN-V	Placebo	Benefit ratio (95% CI)	*p* value
L. crispatus	week 12	≥50% *L. crispatus*	37 (30%)	5 (9%)	3.37 (1.40–8.11)	<0.005
		<50% *L. crispatus*	86 (70%)	51 (91%)		
	week 24	≥50% *L. crispatus*	39 (35%)	4 (8%)	4.49 (1.69–11.90)	–
		<50% *L. crispatus*	74 (65%)	48 (92%)		
Lactobacillus	week 12	≥50% *Lactobacillus*	69 (56%)	27 (48%)	1.16 (0.85–1.59)	–
		<50% *Lactobacillus*	54 (44%)	29 (52%)		
	week 24	≥50% *Lactobacillus*	65 (58%)	17 (33%)	1.76 (1.15–2.68)	–
		<50% *Lactobacillus*	48 (42%)	35 (67%)		

poorly, especially for *Prevotella*-dominated samples (median BC dissimilarity > 0.5), and >40% of samples were almost equally similar to ≥2 CSTs. To identify topics (which can be interpreted as bacterial subcommunities) and estimate topic proportions, we fitted a latent Dirichlet allocation (LDA) model (Bayesian topic model),[53] adapting a previously described approach[50] and constrained topics to consist either of *Lactobacillus* or non-*Lactobacillus* species. We inferred four non-*Lactobacillus* topics (Figures 2A, S1D, and S1E) and defined four *Lactobacillus*-dominated topics: three exclusively comprising a single species (*L. crispatus*, *L. iners*, or *L. jensenii/mulieris*) and one comprising a mixture of remaining *Lactobacillus* species (Figure S1F). Figure 2B displays participant topic proportions throughout the trial.

Three non-*Lactobacillus* topics predominated in both arms pre-MTZ, with *Candidatus* Lachnocurva vaginae (BVAB1), *Gardnerella* (ASV1), or *Prevotella bivia* as predominant taxa, respectively (Figures 2C and S1G). Composition shifted post-MTZ as assessed by Bray-Curtis (BC) dissimilarity (Figure 2D), primarily due to increased *L. iners*, with the most abundant non-*Lactobacillus* topic being *Gardnerella* predominant (Figures 2C and S1G). Another shift occurred between post-MTZ and week 4 visits, greater in the LBP arm (median pairwise BC dissimilarity: 0.81; IQR: 0.50–0.97) than the placebo (median BC dissimilarity: 0.58; IQR 0.41–0.80), driven primarily by increased *L. crispatus* among LBP recipients (Figures 2C, 2D, and S1G). Composition stabilized after week 4 with median BC dissimilarity < 0.55 between all consecutive visits, comparable between arms (Figure 2D). LBP recipients who achieved *L. crispatus* dominance had higher stability (median BC < 0.26) than those who did not (median BC > 0.48, Figure S1H).

Microbiota category analysis also showed stability after week 4. Among LBP recipients with available data, 71% with *L. crispatus* dominance at week 12 or week 24 were also *L. crispatus* dominant at week 4 (Figure 2B; Table S3). Seventy-nine percent of those with *L. crispatus* dominance at week 12 retained it at week 24. However, early *L. crispatus* dominance did not guarantee persistence. Only 49% and 57% of LBP recipients with *L. crispatus* dominance at week 4 retained it at weeks 12 and 24, respectively. By contrast, among 38 LBP recipients with week 4 non-*Lactobacillus* dominance, 79% remained non-*Lactobacillus* dominant at week 12 compared with just 10.5% each for week 4 *L. crispatus* dominance or other *Lactobacillus* dominance (Table S3). Thus, microbiota established early tended to persist through week 24, and early *Lactobacillus* dominance—particularly *L. crispatus* dominance—was strongly linked to ongoing *Lactobacillus* dominance.

CTV-05 was the dominant *L. crispatus* strain in LBP recipients early, but native strains replaced it in a subset of recipients

We next examined what fraction of *L. crispatus* in LBP recipients consisted of CTV-05 versus non-LBP "native" strains, using shotgun metagenomic sequencing (metrics in STAR Methods and Data S2). *L. crispatus* strain genotypes and proportions were inferred from metagenomes using StrainFacts.[54] StrainFacts defines strain genotypes based on single-nucleotide variant (SNV) profiles at defined biallelic sites in the species core genome. In a separate validation study, we found that StrainFacts performs well at inferring LBP (including CTV-05) and native strain genotypes, as well as proportions in samples with ≥5% relative abundance of total *L. crispatus*,[55] so samples were selected for strain analysis based on a 5% threshold. *L. crispatus* strain inference was successful in 313 samples. Based on validation results, a 10% strain fractional abundance was used to define presence in a sample.[55]

We inferred presence of 24 distinct *L. crispatus* strains detected in ≥1 sample within the cohort. To determine which strain represented CTV-05, we generated a closed CTV-05 genome, then determined its StrainFacts genotype. CTV-05's genotype nearly perfectly matched a strain widely prevalent in metagenomes (Jaccard similarity > 0.999), identifying that strain as CTV-05. CTV-05 was inferred in 217 of 256 (84%) of post-randomization LBP recipient samples with ascertainable strain composition but only 5 of 57 (8.7%) samples from placebo recipients or pre-treatment LBP recipients (Figure 3A), suggesting CTV-05 detection was largely accurate but occasionally misclassified native strains as CTV-05. Most participants with an inferred native strain had only one native strain (occasionally two or three) throughout the study, and participants often had the same native strain(s) at multiple visits (Figure 3B). Because strain inference was blinded to participant identity,[54,55] these results support analysis validity.

To experimentally test strain-tracing accuracy, we performed isolations in 12 distinct samples from 6 participants metagenomically predicted to contain no (<0.5%) *L. crispatus* ($n = 2$ samples) or abundant (>40%) *L. crispatus* ($n = 10$). This included samples predicted to contain solely CTV-05, solely native strains, or mixed CTV-05 and native strains (Figures 3C and S2A). We recovered *L. crispatus* from all samples predicted to contain it (106 total

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

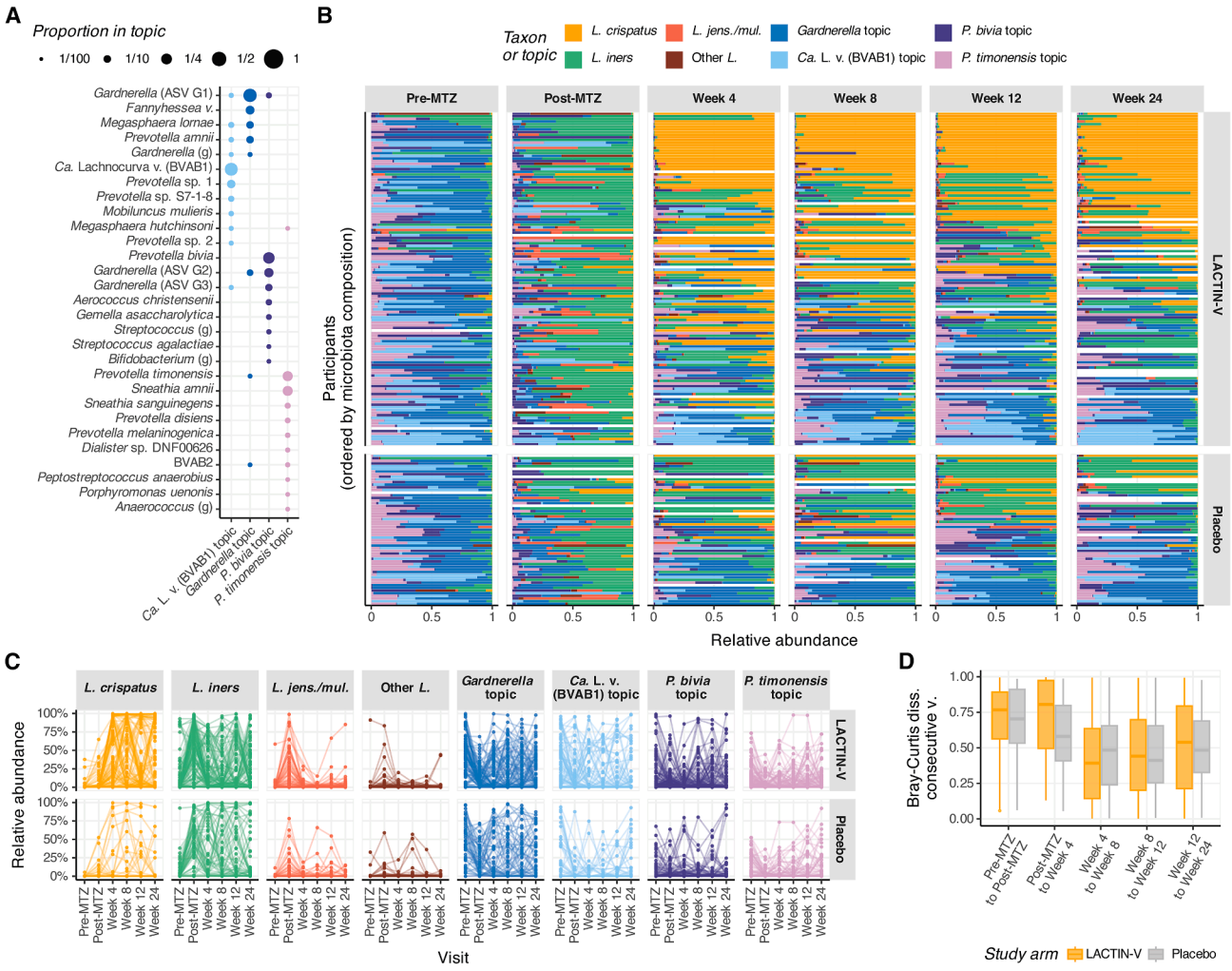


Figure 2. Microbiota composition established early during treatment tended to persist

(A) Taxa proportions in each non-*Lactobacillus* topic for taxa comprising ≥1% of any topic; topic proportions sum to one. Topics are named by predominant taxon. "*Ca.* L. v. (BVAB1)": *Candidatus* Lachnocurva vaginae (BVAB1). "*P. bivia*," *Prevotella bivia*; "*P. timonensis*," *Prevotella timonensis*. See also Figure S1.

(B) Microbiota composition at each scheduled visit for participants with data from ≥3 visits, ordered by microbiota trajectories. See also Table S3.

(C) Relative topic abundance at each visit. Lines connect individual participants' values.

(D) Bray-Curtis (BC) dissimilarity between each participant's microbiota at indicated consecutive visits.

See also Figure S1.

isolates from 10 samples), presumptively identifying isolates by colony morphology and confirmed via genome sequencing (Figures 3D and S2A; Data S3). A phylogenetic tree of *L. crispatus* isolate genomes revealed 7 clades, each containing multiple near-identical genomes. The largest clade, which clustered with CTV-05's genome (Figure 3D, red arrow), included isolates from 7 of 8 samples and all 5 participants with predicted CTV-05 (Figures 3C, 3D, and S2A). Other clades each contained isolates from a single participant. Samples predicted to have two native strains each produced two distinct non-CTV-05 clades, and samples predicted to have a single native strain each produced a single clade. The only exception was a sample (participant STI.00625, week 24) predicted to contain CTV-05 plus a native strain that produced no CTV-05 isolates, but instead produced two non-CTV-05 clades. Among participants with multiple cultured time points, participant STI.00629 produced CTV-05 isolates at week 12 and a native strain at week 24 (consistent with predicted complete strain replacement; Figures 3C, 3D, and S3A), participant STI.00356 produced CTV-05 isolates at week 12 and both CTV-05 and a native strain at week 24 (consistent with predicted partial strain replacement), and participant STI.00625 produced CTV-05 at week 8, a mix of CTV-05 and a native strain at week 12, and a mix of the same native strain plus a new native strain at week 24 (consistent with predictions, except as noted above for week 24). Jaccard similarity between each isolate's StrainFacts genotype and all inferred strains showed each isolate's most closely matching inferred strain was present in the sample from which the isolate was derived (Figures 3C, 3D, and S3B). All members of each clade matched the same inferred strain. In participant STI.00625's week 24

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

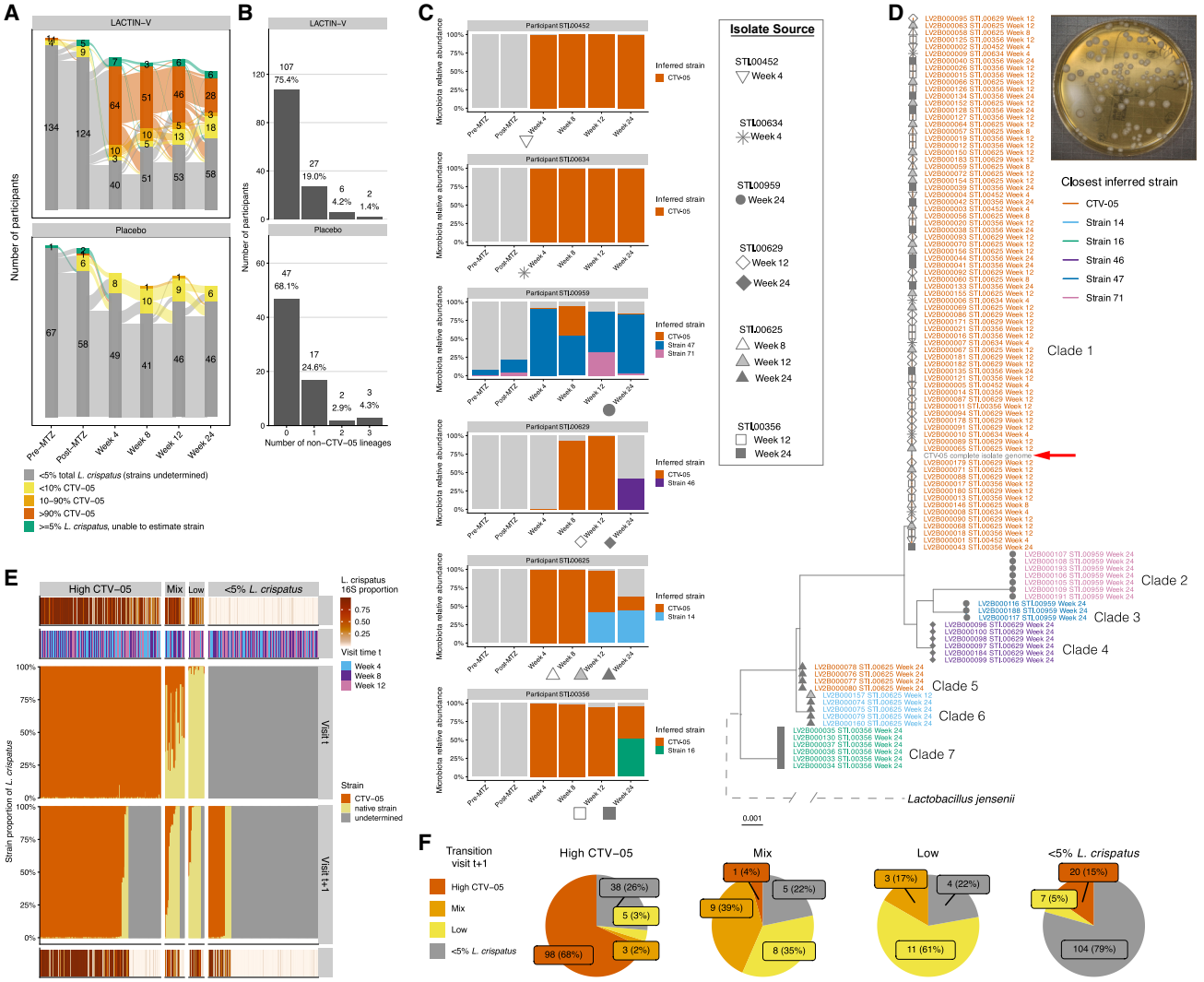


Figure 3. CTV-05 was the dominant *L. crispatus* strain in LBP recipients early, but native strains replaced it in a subset of recipients

(A) CTV-05 and native *L. crispatus* strain dynamics inferred from metagenomes. Samples were categorized by proportional strain abundance: >90% CTV-05 ("high CTV-05"), 10%–90% CTV-05 ("mixed"), and <10% CTV-05 ("low CTV-05"). Green: samples with technical failures in strain inference, see also Data S2.

(B) Number (percentage) of participants in each arm who had the indicated number of native strains detected at least once during the trial.

(C) Metagenomically inferred *L. crispatus* strain composition of LBP recipients selected for isolations ($n = 6$). Color indicates unique inferred strains; gray shows summed abundance of non-*L. crispatus* taxa. Symbols indicate samples used for *L. crispatus* isolation. See also Figure S2.

(D) Phylogenetic tree of CTV-05 (arrow) and *L. crispatus* isolates from (C). Symbols indicate isolate source; colors show each isolate's closest matching inferred strain (see C). Inset: typical *L. crispatus* colony morphology. See also Figure S2 and Data S3.

(E) Transition plot for LBP recipients showing consecutively scheduled pairs of post-randomization visits with *L. crispatus* strain data or with insufficient *L. crispatus* for strain inference. Starting ("visit t," top row) and subsequent ("visit t + 1," bottom row) *L. crispatus* relative abundance. Second row: study week at visit t. Third and fourth rows: CTV-05 and summed native strain proportions at visit t (third row) and visit t + 1 (fourth row). Visit pairs are grouped by visit t strain category and ordered by visit t + 1 CTV-05 proportion.

(F) Pie charts for visit t strain categories (E), summarizing strain category frequencies at visit t + 1.

See also Figure S2.

sample, the closest inferred strain for one clade's genomes (clade 5) was CTV-05 but at lower similarity than the true CTV-05 clade (Figure S3B). Thus, isolations showed that strain inference was accurate but occasionally misclassified native strains as CTV-05.

To analyze *L. crispatus* strain dynamics, we classified samples into three categories: >90% inferred CTV-05 fractional strain abundance ("high-CTV-05"), 10%–90% CTV-05 ("mixed strains"), and <10% CTV-05 ("low-CTV-05" or high native strain(s); Figure 3A). From week 4 onward, most LBP recipients with analyzable *L. crispatus* had high-CTV-05 strain abundance, but the number and proportion with high native strains progressively increased, while numbers and proportions with high-CTV-05 or mixed strains decreased (Figure 3A). At week 4, 96.1% ($n = 74$) of LBP recipients

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

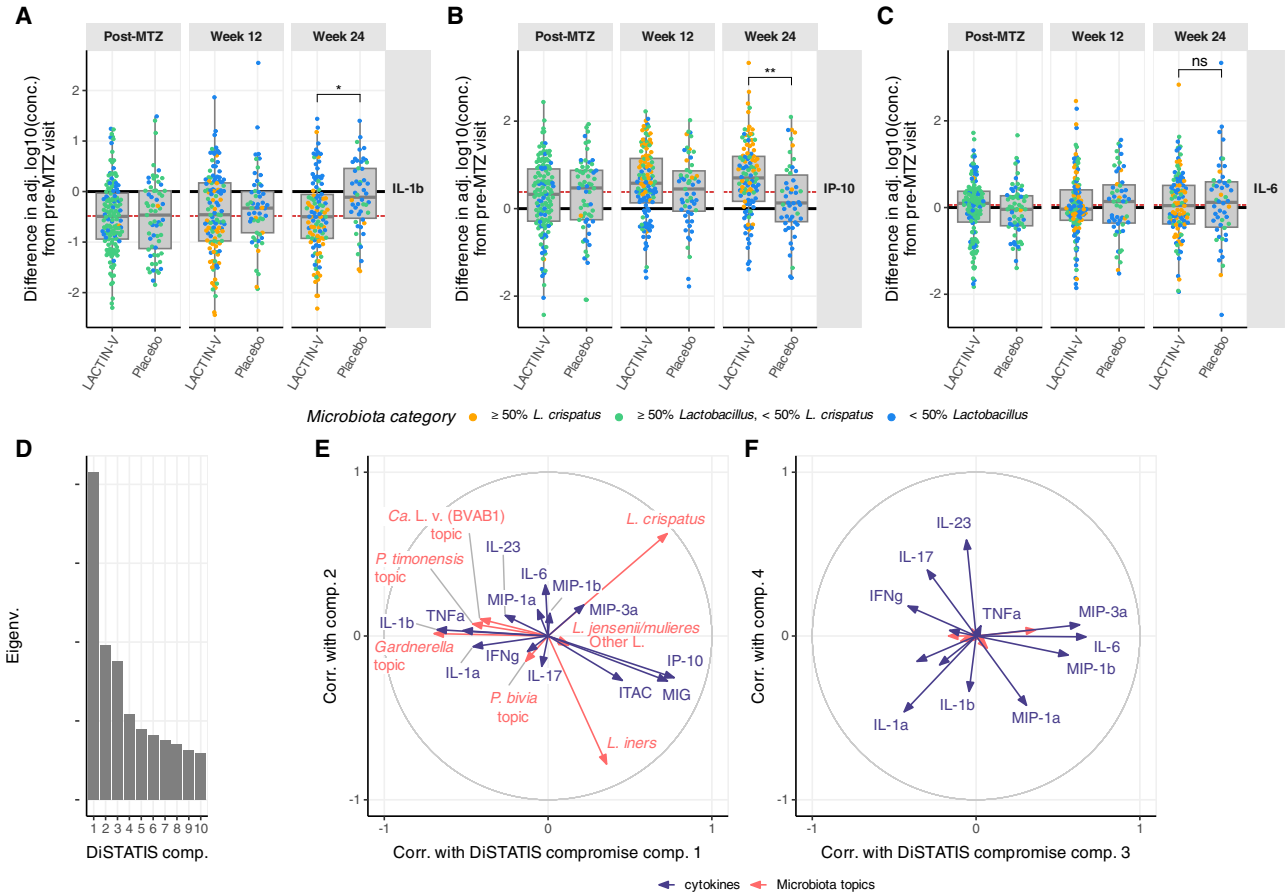


Figure 4. MTZ reduced mucosal inflammation associated with low *Lactobacillus* abundance, which reverted by week 24 in placebo but not LBP recipients

(A) Differences in log10-transformed, adjusted IL-1β concentrations between participants' pre-MTZ and week 12 (left) or week 24 (right) visits. Significance of week 24 differences is shown (Wilcoxon rank-sum test with Benjamini-Hochberg correction, $^{**}p < 0.01$, $^{*}p < 0.05$, ns: $p > 0.05$). See also Figure S3.

(B) Same as (A) for IP-10.

(C) Same as (A) and (B) for IL-6.

(D) DiSTATIS scree plot: eigenvalues (a.u.) of DiSTATIS compromise for first 10 latent components.

(E) DiSTATIS correlations between the 1st and 2nd compromise latent dimensions and microbiota topic proportions or cytokine-transformed concentrations.

(F) Same as (E) for 3rd and 4th latent components.

See also Figure S3.

with ascertainable strains had high-CTV-05 or mixed strains, while just 3.9% ($n = 3$) had high native strains. However, by week 24, 63.2% ($n = 31$) had high-CTV-05 or mixed strains, while 36.8% ($n = 18$) had high native strains. Participant-level analysis showed that, among 144 visits with high-CTV-05, two main outcomes were observed at the next visit: 68% ($n = 98$) maintained high-CTV-05 and 26% ($n = 38$) fell below the strain detection threshold (Figures 3E and 3F). Among 18 visits with high native strains, 61% ($n = 11$) retained high native strains, and none transitioned to high-CTV-05. Among 23 visits with mixed strains, just one (4%) transitioned to high-CTV-05, while 39% ($n = 9$) remained mixed, and 35% ($n = 8$) transitioned to high native strains, including participants still receiving LACTIN-V treatment (Figure S2C). Thus, CTV-05 was frequently the primary *L. crispatus* strain in LBP recipients, but when native strains dominated or co-occurred with CTV-05, they often subsequently replaced it, whereas CTV-05 almost never replaced native strains.

MTZ reduced mucosal inflammation associated with low *Lactobacillus* abundance, which reverted by week 24 in placebo but not LBP recipients

To assess treatment effects on vaginal inflammation, we measured cytokines and chemokines in vaginal swab eluates (Figure S3A).[6,50] Analyte concentrations were strongly positively correlated (Figure S3B), suggesting a technical "size-effect"[56] from sampling variation, for which we adjusted using principal component 1 (PC1) subtraction. Participant-level immune dynamics were assessed by comparing each participant's adjusted pre-MTZ concentration for each analyte with concentrations at subsequent visits. Interleukin (IL)-1β decreased post-MTZ, whereas interferon gamma-induced protein 10 (IP-10) rose, consistent with reported microbiota and BV associations[5,6,57,58] (Figures 4A and 4B). Shifts persisted in both arms through week 12, but placebo recipients returned to pre-MTZ levels by week 24, producing differences of −0.39 (95% CI: −0.66 to 0.14)

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

between LBP and placebo for IL-1β and 0.51 (95% CI: 0.21–0.79) for IP-10. By contrast, IL-6—which has been reported not to differ in BV[57]—remained unchanged throughout (Figure 4C).

Week 24 results for IL-1β and IP-10 suggested associations with total *Lactobacillus* relative abundance, which differed between arms at that time point (Table 1). Microbiota category analysis confirmed that lower IL-1β and higher IP-10 were associated with *Lactobacillus* dominance (Figures 4A and 4B). Overall microbiota composition (topic proportions) significantly correlated with concentrations of the 13 analyzed cytokines/chemokines (RV coefficient 0.18; p value < 0.005). Analyzing visits individually produced similar results (Figure S3C), except for a lower correlation at the post-MTZ visit (where bacterial load varied widely). To characterize microbiota-immune relationships in greater detail, we employed DiSTATIS,[59,60] which infers a consensus representation of the two datasets from between-sample dissimilarities. Projecting microbiota topics and cytokines/chemokines revealed that the first latent dimension discriminated *Lactobacillus*-dominated from non-*Lactobacillus*-dominated samples (Figures 4D–4F). Monokine induced by interferon gamma (MIG), IP-10, and interferon-inducible T cell alpha chemoattractant (ITAC) were positively correlated (covaried with *Lactobacillus*), while IL-1β, tumor necrosis factor α (TNF-α), and IL-1α were negatively correlated (covaried with non-*Lactobacillus* topics; Figure 4E). The second dimension discriminated *L. crispatus* from *L. iners*, but most cytokines and chemokines showed little correlation, indicating minimal immune association with individual *Lactobacillus* species. Most cytokines and chemokines had high correlations with the 3rd or 4th latent components, but microbiota topics did not, indicating a degree of microbiota-independent immune variation (Figure 4F). In these dimensions, interferon gamma (IFN-γ) and IL-17 correlated, but they were anti-correlated with IL-6, macrophage inflammatory protein (MIP)-3α, MIP-1β, and MIP-1α. Analysis distinguishing CTV-05 from other *L. crispatus* strains showed no clear differences in microbiota-immune correlations (Figures S3D–S3G).

LBP treatment benefits differed for women with different pre-MTZ microbiota

We performed exploratory post hoc analyses assessing whether participants with different pre-MTZ microbiota composition responded differently to LBP (treatment effect heterogeneity). Stratifying participants by their most abundant pre-MTZ taxon (genus-level) identified four groups: *Lactobacillus* predominant, *Gardnerella* predominant, *Prevotella* predominant, and *Ca.* Lachnocurva vaginae (BVAB1) predominant. LBP benefit was evaluated for two separate outcomes: *L. crispatus* dominance (at week 12 or 24) and BV recurrence (by week 12 or 24) (Figures 5A and 5B). Analysis of *L. crispatus* dominance suggested that some groups differentially benefited from LBP treatment (adjusted p values = 0.11 and 0.03 at weeks 12 and 24, respectively; Figures 5C and 5D). LBP recipients with pre-MTZ *Lactobacillus*, *Gardnerella*, or *Prevotella* predominance attained higher rates of subsequent *L. crispatus* dominance than their placebo counterparts, with modestly larger differences at week 24 than at week 12 (Figure 5D). In contrast, LBP treatment was not associated with higher rates of *L. crispatus* dominance for participants with pre-MTZ *Ca.* Lachnocurva vaginae (BVAB1) predominance. LBP benefit in preventing rBV also differed by pre-MTZ microbiota (adjusted p values < 0.05 and < 0.01 at weeks 12 and 24; Figures 5E and 5F). BV recurred in almost all placebo recipients with pre-MTZ *Prevotella* predominance by week 24 versus just over half of LBP recipients. By contrast, LBP recipients with pre-MTZ *Ca.* Lachnocurva vaginae (BVAB1) predominance had slightly higher rBV rates than placebo (Figures 5E and 5F). Sensitivity analyses grouping participants by pre-MTZ CST[52] (Figure S4) or using a model-based approach relying on topic relative abundances provided largely consistent results. Collectively, these exploratory analyses suggested a lack of LBP benefit for participants with pre-MTZ *Ca.* Lachnocurva vaginae (BVAB1) predominance and variable degrees of benefit for all others.

Microbiota, vaginal, immune, and sociodemographic factors were associated with *L. crispatus* dominance in LBP recipients

To identify factors associated with achieving *L. crispatus* dominance in LBP recipients, we performed multiblock partial least-squares discriminant analyses (MB-PLS-DA) on LACTIN-V arm data (Figures 6A–6C and S5), leveraging the biological data described above and sociodemographic, clinical, and behavioral data collected during the clinical trial (variables and their unadjusted relationships to rBV and microbiota outcomes in Data S4). MB-PLS-DA simultaneously examines how variables contribute to a model's ability to explain outcomes both individually ("variable importance") and in thematic groups of variables ("block importance").[61,62] A block could thus contribute to model performance either through modest contributions from each of its variables or through strong contributions from a few variables. We defined three models corresponding to different trial phases: "initial phase" (post-MTZ to week 4), "continuation phase" (weeks 4–12), and "follow-up phase" (weeks 12–24, Figure S5A). Explanatory variables were grouped into 13 total blocks, characterizing participant demographics, baseline (pre-MTZ) vaginal ecosystem, vaginal ecosystem at the preceding visit, and clinical/behavioral parameters (Figure 6A; Data S4). For each phase, we evaluated whether models could discriminate between three microbiota categories: ≥50% *L. crispatus*, ≥50% *Lactobacillus* but <50 % *L. crispatus*, or <50% *Lactobacillus*. We used cross-validation to select latent component numbers and avoid over-fitting, and nested models to further assess block contributions (Figures S5A–S5C). Results showed that initial-phase colonization was more difficult to explain, as well as being explained by more factors, than colonization at later phases (see STAR Methods; Figures 6A, S5B, and S5C).

In the initial-phase model for the LBP arm, the most important block was previous (i.e., post-MTZ) vaginal environment (Figure 6A, left), which included post-MTZ total bacterial load, α-diversity, and pH. Lower values were associated with a higher chance of week 4 *L. crispatus* dominance, and these were among the 5 most important variables (Figures 6B, 6C, and S5D). The next most important blocks were demographics and blocks characterizing pre-MTZ vaginal environment, cytokine concentrations, and microbiota composition (Figures 6A and S5D). Specifically, high pre-MTZ α-diversity and pH were negatively associated with week 4 *L. crispatus* dominance (Figure 6B). Interestingly, pre-MTZ IL-1β levels were positively associated

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

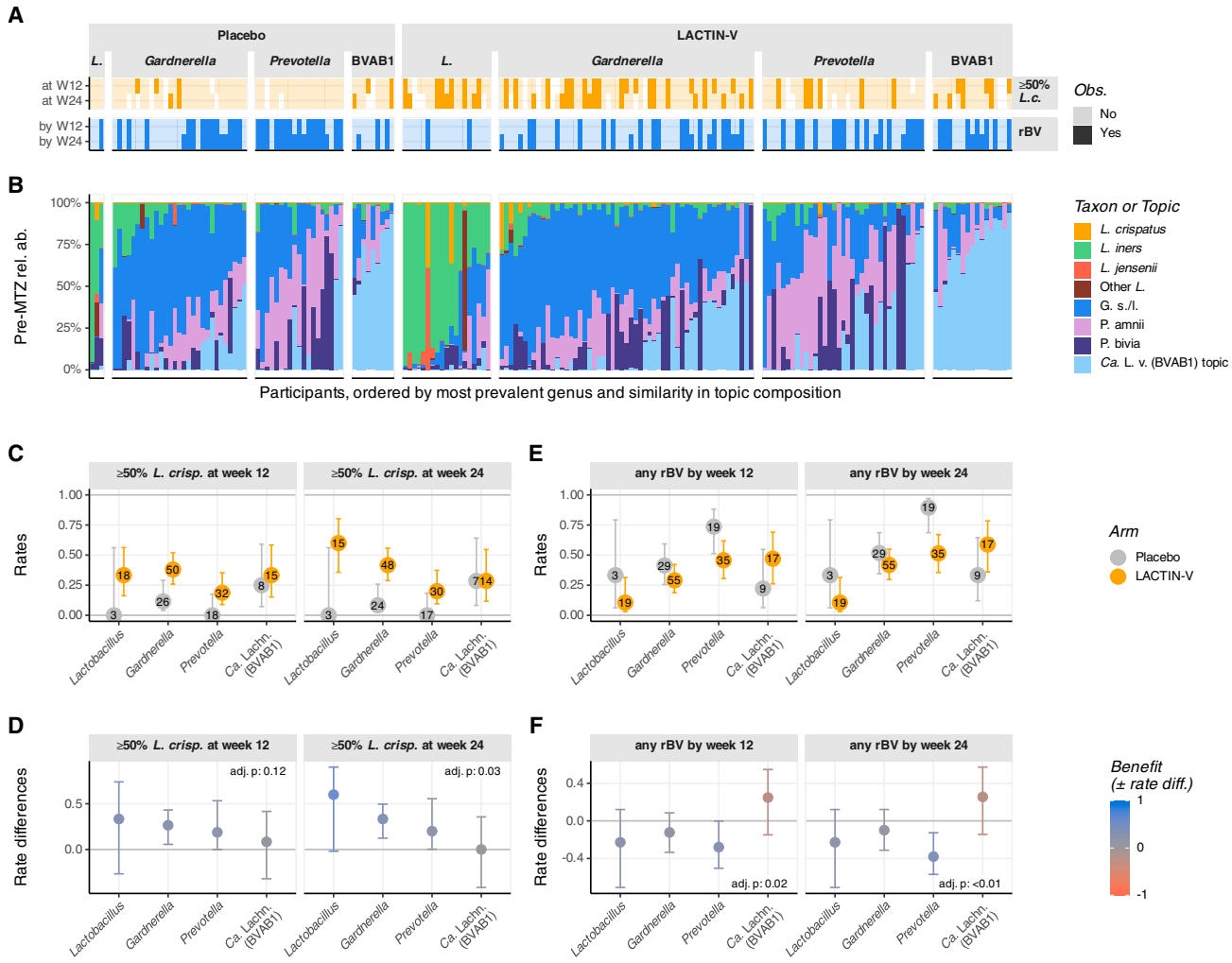


Figure 5. LBP treatment benefits differed for women with different pre-MTZ microbiota

(A) Weeks 12 and 24 outcomes, including *L. crispatus* dominance (dark orange), <50% *L. crispatus* (light orange), rBV (dark blue), or no rBV (light blue), grouped horizontally by predominant pre-MTZ genus ("L.," *Lactobacillus*; "BVAB1," *Ca.* Lachnocurva vaginae), showing groups with ≥2 participants per arm (white: no data).

(B) Pre-MTZ microbiota composition organized as in (A).

(C) Rates and 95% CI of *L. crispatus* dominance at week 12 (left) or 24 (right) by pre-MTZ microbiota group (A).

(D) Between-arm rate differences (LBP − placebo) and 95% CI based on (C). Color shows degree of benefit in achieving *L. crispatus* dominance (blue: benefit with LBP). Significance was determined by analysis of deviance test, with Benjamini-Hochberg adjustment for *p* values across (D) and (F).

(E) Same as (C), but for rBV by week 12 or 24.

(F) Same as (D), but for rBV by week 12 or 24 (blue: LBP benefit in reducing rBV).

See also Figure S4.

with week 4 *L. crispatus* dominance, while pre-MTZ IP-10 and MIG levels showed the opposite association (Figures 6B, 6D, and S5D). Self-identified race and education were the most important demographic variables and slightly contributed to model performance, with participants identifying as more highly educated and white achieving higher rates of week 4 *L. crispatus* dominance (Figure S5D). However, self-identified race and education were highly correlated with each other (χ^2 *p* value < 0.01; Figure S5E) and with certain pre-MTZ vaginal characteristics such as microbiota α-diversity as well as study site (Figures S5F and S5G), but adding study site as a demographic block variable did not improve model performance. Thus, self-identified race—itself a non-biological, socially defined category[63,64]—was confounded with measured and unmeasured socioeconomic, geographic, and clinical factors in this cohort, such that individual relationships with *L. crispatus* colonization could not be identified. Notably, women self-identifying as white had higher rates of *L. crispatus* dominance in both the placebo and LBP arms, but LBP treatment increased *L. crispatus* dominance relative to placebo to a similar degree for all self-declared racial groups at week 12 (Figures 6E and 6F). There was a modest trend toward greater benefit in women identifying as white at week 24, which was not statistically significant (adjusted *p* values = 0.57 at both weeks). Remaining blocks were not

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

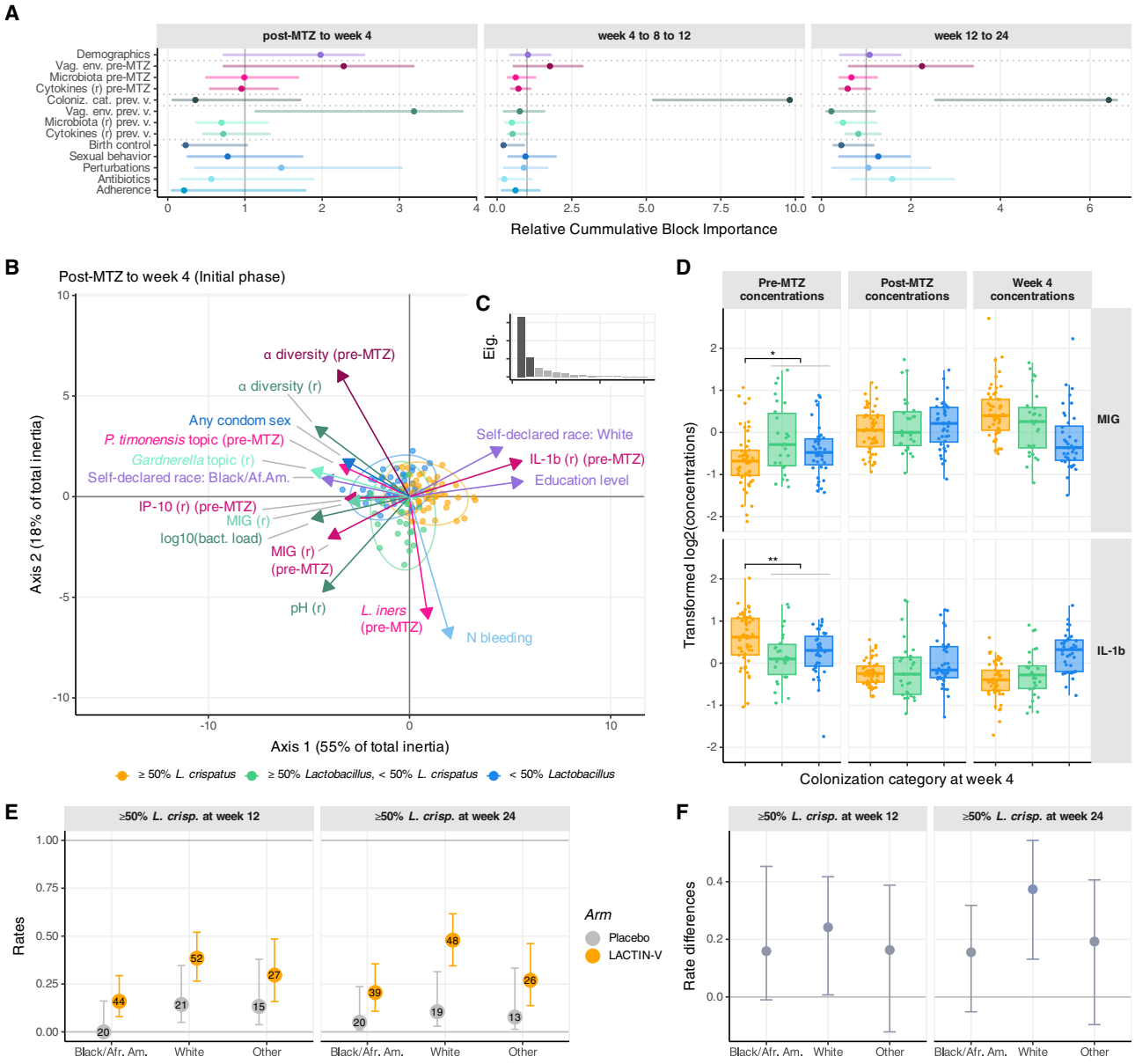


Figure 6. Microbiota, vaginal, immune, and sociodemographic factors were associated with *L. crispatus* dominance in LBP recipients

(A) Relative cumulative block importances from MB-PLS-DA analysis of variables predictive of microbiota categories defined in Figure 1B. Point estimates and 95% CI are shown for each variable block for initial-phase (left), continuation-phase (middle), and follow-up-phase models (right) for the LBP arm. Black vertical lines show the (null) hypothesis that all blocks have equal importance. See also Figure S5 and Data S4.

(B) Bi-plot of the initial-phase model (LBP arm). Dots (colors: week 4 microbiota category) represent participant scores; arrows (colors: blocks as in A) show loadings of the most important variables in the first 2 latent variables. Perfect separation of microbiota categories would indicate perfect model performance. See also Figure S5.

(C) Scree plot of first 15 eigenvalues of model from (B). Dark gray: first two latent variables.

(D) Concentrations of cytokines (MIG and IL-1β) whose pre-MTZ levels (left) predicted week 4 category in the initial-phase model (LBP arm). Post-MTZ (middle) and week 4 (left) concentrations are also shown. Data are grouped and colored by week 4 microbiota category in all panels. Left panel shows *t* tests comparing pre-MTZ concentrations for participants achieving *L. crispatus* dominance at week 4 to all others (**$p < 0.01$, *$p < 0.05$). See also Figures S6 and S7.

(E) Rates and 95% CI of *L. crispatus* dominance at week 12 (left) or 24 (right) by self-declared race.

(F) Between-arm rate differences (LBP − placebo) and 95% CI for achieving *L. crispatus* dominance based on (E).

important for initial-phase predictions, although within these blocks, factors with higher variability across participants, such as menstrual bleeding and sexual activity, showed greater variable importance (Figures 6A, 6B, and S5D).

In the LBP arm continuation-phase (weeks 4–12), microbiota category at the previous visit was by far the most important block (Figure 6A). Although no other blocks significantly improved model performance (Figure S5C), individual variables, including

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

sexual activity, bleeding, douching, and non-hormonal IUDs, were negatively associated with *L. crispatus* dominance (Figure S5D). In the follow-up phase (weeks 12–24), microbiota category at the previous visit again had high importance, and sexual behavior and antibiotic use mildly contributed to explaining outcomes (Figures 6A, right and S5D).

Analogous placebo arm models showed no significant predictive value for the initial phase or follow-up phase in cross-validation (Figures S6 and S7). The placebo continuation-phase model had modest predictive value, with the previous-visit microbiota category serving as the best predictor (Figures S6 and S7).

DISCUSSION

Vaginal LBPs offer significant promise to improve health, but mechanisms and strain dynamics of LBP colonization and efficacy are incompletely understood.[18,46] We analyzed samples and data from a randomized, placebo-controlled trial of LACTIN-V—a single-strain *L. crispatus* LBP to prevent BV recurrence[37]—employing microbiome sequencing, immunologic characterization, and multi-omic analyses to investigate LBP effects and correlates of success. LBP treatment produced *L. crispatus*-dominant vaginal microbiota in 30% and 35% of recipients at weeks 12 and 24, respectively, substantially superior to placebo. *L. crispatus* colonization was primarily due to the LACTIN-V strain CTV-05, although CTV-05 diminished and native *L. crispatus* strains increased over time. *Lactobacillus* dominance—particularly *L. crispatus* dominance—at early time points was linked to persistent *Lactobacillus* dominance, while early non-*Lactobacillus* dominance also tended to persist. Exploratory analyses showed that LBP benefits differed by baseline microbiota composition and identified microbial, immune, demographic, and behavioral factors associated with *L. crispatus* dominance among LBP recipients.

Investigation of *L. crispatus* strain dynamics showed that CTV-05 accounted for most *L. crispatus* colonization in LBP recipients, but the number and proportion with high native strains progressively increased over time, from 3.9% at week 4 to 36.8% at week 24. When CTV-05 co-occurred with native strains, the native strains frequently replaced CTV-05 at subsequent visits. By contrast, native strain replacement by CTV-05 was very rare. Further research is needed to determine whether native strains that outcompete CTV-05 have functional characteristics providing selective advantages or whether initial CTV-05 colonization establishes a permissive environment for native strains. It is also unknown whether strain dynamics observed in this US cohort will be replicated in non-US populations treated with LACTIN-V or in women receiving LBPs containing *L. crispatus* from different clinical or geographic sources.

Vaginal mucosal cytokine and chemokine levels changed significantly with microbiota shifts from BV to *Lactobacillus dominance*, and reverted if microbiota returned to non-*Lactobacillus* dominance. Specifically, cytokines including IL-1α, IL-1β, and TNF-α increased with decreasing *Lactobacillus*, whereas IP-10, MIG, and ITAC showed the opposite patterns, consistent with prior reports.[5,6,57,58] A prior study of a small subset of trial participants reported lower IP-10 levels at study end in women with CTV-05 compared with other *L. crispatus*,[65] but we saw no similar association in the larger cohort. Favorable cytokine changes persisted in LBP recipients at week 24 compared with placebo, largely due to higher rates of overall *Lactobacillus* dominance rather than CTV-05-specific effects.

LACTIN-V efficacy varied with differences in pre-MTZ microbiota composition. Compared with placebo, LBP treatment particularly reduced rBV among participants with pre-MTZ *Prevotella* predominance, whereas participants with pre-MTZ *Ca.* Lachnocurva vaginae (BVAB1) predominance uniquely trended toward more rBV with the LBP and showed little or no benefit in achieving *L. crispatus* dominance. Reasons for these patterns are unclear, but they may reflect greater competitive ability or MTZ resistance in *Ca.* Lachnocurva vaginae or its co-occurring species.[66,67] Possible differences in *Ca.* Lachnocurva vaginae response to oral versus intravaginal MTZ have been reported.[68] *Ca.* Lachnocurva vaginae remains uncultured,[69,70] so our results emphasize that cultivating and phenotypically characterizing it are key research priorities. Importantly, these exploratory post hoc analysis results should be considered hypothesis-generating observations requiring validation in future studies.

Integrated multiblock analysis of LBP recipients showed that low bacterial load, vaginal pH, and α-diversity at the post-MTZ visit were associated with subsequent *L. crispatus* dominance, elucidating the microbiota dynamics underlying a post hoc analysis that found that participants clinically cured of BV at the post-MTZ visit had less subsequent rBV.[18] Low pre-MTZ MIG and IP-10 levels and high pre-MTZ IL-1β levels were also associated with early *L. crispatus* dominance—an unexpected finding given their opposite correlations with *Lactobacillus* at concurrent visits. Reasons are unclear but may involve a more vigorous pre-MTZ immune response helping clear BV-associated bacteria, the presence of inflammatory but MTZ-responsive bacteria, or other mechanisms. After week 4, the strongest predictor of microbiota category was the prior visit's microbiota, indicating that early post-LBP microbiota best predicts subsequent outcomes.

Self-declared race and education were also linked to microbiota composition at week 4 and beyond in LBP recipients, but LBP treatment increased *L. crispatus* dominance compared with placebo for all racial groups. Prior US studies have linked BV and non-*Lactobacillus* dominance to self-identified non-white race and/or Hispanic ethnicity.[12,51] However, education and race correlated closely in our cohort, differed by study site, and thus likely correlated with unmeasured factors relevant to microbiota composition, including nutrition and health service access, environmental exposure, and psychosocial stressors that elevate levels of hormones such as cortisol, which can alter sex hormone regulation, menstrual bleeding, and cervicovaginal epithelial function.[63,64,71,72] BV's complex relationship with socially defined categories like race/ethnicity is highlighted by the fact that BV disparities are observed in ancestrally unlinked, socioeconomically disadvantaged populations globally, including nomadic women in Iran,[73] aboriginal women in Canada,[74] and Tibetan women in China,[75] while lower educational and socioeconomic status correlated with lower *L. crispatus* dominance in a homogeneous cohort of Caucasian Finnish women.[76] A recent Chinese study reported associations between vaginal microbiota and environmental factors, including ambient temperature and pollutant exposure, neither of which were addressable in our

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

study.[77] Thus, our results should not be interpreted as showing that education or race are mechanistic causes of BV risk or treatment outcome, and we believe that additional research on determinants of these associations is needed. Compared with placebo, LBP treatment increased *L. crispatus* dominance for women identifying as black to a similar degree as other groups, even though their absolute rates of *L. crispatus* dominance were lower in both the placebo and LBP arms. Thus, further improving LBPs or other BV therapies has the potential to offer particular benefit for these women.

Despite evidence that sexual activity influences BV risk and vaginal microbiota composition,[78] self-reported sexual behavior was not strongly predictive of microbiota composition in LBP recipients. This may be due to consistent sexual behavior throughout the trial, such that the effects of sex occurred early and were obscured by other factors associated with early *L. crispatus* dominance. Because self-report is often an unreliable measure of true sexual activity,[79–82] inaccurate self-reporting may also have decreased the power to detect these effects. We stress that these results are exploratory and require future validation.

In summary, our analysis of microbiota and strain dynamics underlying the clinical effects of a vaginal *L. crispatus* LBP identifies key microbiota and host factors associated with treatment outcomes. If validated, these findings may help identify patients likely to benefit from treatment and guide discovery of novel targets to enhance LBP efficacy and improve women's health globally.

Limitations

This study has several limitations. First, samples from a few participants were unavailable (10 of 152 LBP and 5 of 76 placebo recipients), and technical sequencing failures affected four samples, so our cohort was slightly smaller than the original trial. Second, the study was powered to detect differences in rBV, so its limited size reduced the power to detect heterogeneity in treatment effects or correlates of *L. crispatus* dominance. Third, the sampling schedule did not permit high-resolution temporal assessment of microbiota dynamics. Fourth, participants were followed for 24 weeks post-randomization (including 13 weeks post-treatment), precluding assessment of longer-term microbiota patterns. Finally, detailed sexual, contraceptive, hygiene, and clinical data were available, but only a narrow range of environmental and sociodemographic variables were collected, limiting our analysis of factors that might help explain associations with microbiota composition.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Douglas S. Kwon (dkwon@mgh.harvard.edu).

Materials availability

There are restrictions on the availability of *L. crispatus* strain CTV-05, including CTV-05 strains re-isolated from trial samples as part of this study, because of intellectual property and commercial interests of Osel, Inc., the manufacturer of LACTIN-V, an investigational drug under Investigational New Drug (IND) status with the US Food and Drug Administration. Requests to use CTV-05 for academic research purposes under a materials transfer agreement should be directed to Osel, Inc., at tparks@oselinc.com or info@oselinc.com.

All other unique resources generated in this study are available from the lead contact with a completed materials transfer agreement.

Data and code availability

- Sequence data and genome assemblies have been deposited under NCBI BioProject: PRJNA1303956 and are publicly available as of the date of publication.
- Clinical, processed, and transformed omics data are available on Zenodo: https://doi.org/10.5281/zenodo.17755173.
- Original code is available on Zenodo: https://doi.org/10.5281/zenodo.18524760.
- Any additional information required to reanalyze the data reported in this work is available from the lead contact upon request.

ACKNOWLEDGMENTS

We thank members of the Gates Foundation Vaginal Microbiome Research Consortium for helpful discussions. This work was supported by a grant from the Gates Foundation (INV-033690) to D.S.K. S.M.B. was partially supported by NIH grant 1K08AI171166.

AUTHOR CONTRIBUTIONS

D.S.K., S.P.H., S.M.B., and L.S. conceptualized the research, with input from J.E., C.M.M., A.H., and C.R.C.; D.S.K. and S.P.H. supervised the project; S.M.B., L.S., J.E., J.X., S.H., J.S., and S.P.H. developed and validated methods and models; S.M.B., J.X., S.H., A.S., C.M.M., T.P.P., A.K., and F.A.H. generated samples and performed experiments; L.S., S.M.B., J.E., J.S., A.H., A.K., F.A.H., and S.P.H. performed data analysis; L.S., S.M.B., and D.S.K. led the writing of the manuscript, with contributions from S.P.H., J.E., J.X., S.H., C.M.M., A.H., T.P.P., A.K., F.A.H., and C.R.C.; all authors provided critical feedback on the methods, analysis, and writing.

DECLARATION OF INTERESTS

C.M.M. has been a consultant for Freya Biosciences and serves on the scientific advisory board of Ancilia Biosciences and Concerto Biosciences. C.M.M. has a financial interest in Ancilia Biosciences, a company developing a new class of live biotherapeutics and other bacterial products. C.M.M.'s interests were reviewed and are managed by MGH and Mass General Brigham in accordance with their conflict-of-interest policies. C.R.C. has served as a scientific advisor for Osel, Inc. and Evvy and has stock options from both. The UCSF Conflict of Interest Committee approved a plan to minimize his potential conflict of interest. T.P.P. is an employee of Osel, Inc. and an inventor on US Patent No. 11,083,761, European Patent Application No. 18827562.2, US Patent Application No. 18/277,762, and European Patent Application No. 22707210.5, related to the formulation and application of LACTIN-V.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Study design and sample collection
  - Bacterial isolations and cultivation
- METHOD DETAILS
  - DNA extraction for short-read sequencing
  - Bacterial 16S rRNA gene sequencing
  - Bacterial short-read genome sequencing
  - Bacterial 16S rRNA gene sequence annotation
  - Total bacterial load via qPCR
  - Assessing baseline balance between arms
  - Calculating benefit ratios

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.chom.2026.03.003.

Received: August 31, 2025
Revised: December 24, 2025
Accepted: March 2, 2026

REFERENCES

1. Peebles, K., Velloza, J., Balkus, J.E., McClelland, R.S., and Barnabas, R.V. (2019). High Global Burden and Costs of Bacterial Vaginosis: A Systematic Review and Meta-Analysis. Sex. Transm. Dis. *46*, 304–311. https://doi.org/10.1097/OLQ.0000000000000972.
2. Workowski, K.A., Bachmann, L.H., Chan, P.A., Johnston, C.M., Muzny, C.A., Park, I., Reno, H., Zenilman, J.M., and Bolan, G.A. (2021). Sexually Transmitted Infections Treatment Guidelines, 2021. MMWR Recomm. Rep. *70*, 1–187. https://doi.org/10.15585/mmwr.rr7004a1.
3. Brusselmans, J., De Sutter, A., Devleesschauwer, B., Verstraelen, H., and Cools, P. (2023). Scoping review of the association between bacterial vaginosis and emotional, sexual and social health. BMC Womens Health *23*, 168. https://doi.org/10.1186/s12905-023-02260-z.
4. Bilardi, J.E., Walker, S., Temple-Smith, M., McNair, R., Mooney-Somers, J., Bellhouse, C., Fairley, C.K., Chen, M.Y., and Bradshaw, C. (2013). The burden of bacterial vaginosis: women's experience of the physical, emotional, sexual and social impact of living with recurrent bacterial vaginosis. PLoS One *8*, e74378. https://doi.org/10.1371/journal.pone.0074378.
5. Anahtar, M.N., Byrne, E.H., Doherty, K.E., Bowman, B.A., Yamamoto, H.S., Soumillon, M., Padavattan, N., Ismail, N., Moodley, A., Sabatini, M.E., et al. (2015). Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. Immunity *42*, 965–976. https://doi.org/10.1016/j.immuni.2015.04.019.
6. Gosmann, C., Anahtar, M.N., Handley, S.A., Farcasanu, M., Abu-Ali, G., Bowman, B.A., Padavattan, N., Desai, C., Droit, L., Moodley, A., et al. (2017). Lactobacillus-Deficient Cervicovaginal Bacterial Communities Are Associated with Increased HIV Acquisition in Young South African Women. Immunity *46*, 29–37. https://doi.org/10.1016/j.immuni.2016.12.013.
7. Atashili, J., Poole, C., Ndumbe, P.M., Adimora, A.A., and Smith, J.S. (2008). Bacterial vaginosis and HIV acquisition: a meta-analysis of published studies. AIDS *22*, 1493–1501. https://doi.org/10.1097/QAD.0b013e3283021a37.
8. Hillier, S.L., Martius, J., Krohn, M., Kiviat, N., Holmes, K.K., and Eschenbach, D.A. (1988). A case-control study of chorioamnionic infection and histologic chorioamnionitis in prematurity. N. Engl. J. Med. *319*, 972–978. https://doi.org/10.1056/NEJM198810133191503.
9. Leitich, H., and Kiss, H. (2007). Asymptomatic bacterial vaginosis and intermediate flora as risk factors for adverse pregnancy outcome. Best Pract. Res. Clin. Obstet. Gynaecol. *21*, 375–390. https://doi.org/10.1016/j.bpobgyn.2006.12.005.
10. Brusselaers, N., Shrestha, S., van de Wijgert, J., and Verstraelen, H. (2019). Vaginal dysbiosis and the risk of human papillomavirus and cervical cancer: systematic review and meta-analysis. Am. J. Obstet. Gynecol. *221*, 9–18.e8. https://doi.org/10.1016/j.ajog.2018.12.011.
11. Allsworth, J.E., and Peipert, J.F. (2007). Prevalence of bacterial vaginosis: 2001–2004 National Health and Nutrition Examination Survey data. Obstet. Gynecol. *109*, 114–120. https://doi.org/10.1097/01.AOG.0000247627.84791.91.
12. Kenyon, C., Colebunders, R., and Crucitti, T. (2013). The global epidemiology of bacterial vaginosis: a systematic review. Am. J. Obstet. Gynecol. *209*, 505–523. https://doi.org/10.1016/j.ajog.2013.05.006.
13. Marconi, C., Duarte, M.T.C., Silva, D.C., and Silva, M.G. (2015). Prevalence of and risk factors for bacterial vaginosis among women of reproductive age attending cervical screening in southeastern Brazil. Int. J. Gynaecol. Obstet. *131*, 137–141. https://doi.org/10.1016/j.ijgo.2015.05.016.
14. Bradshaw, C.S., and Sobel, J.D. (2016). Current Treatment of Bacterial Vaginosis—Limitations and Need for Innovation. J. Infect. Dis. *214*, S14–S20. https://doi.org/10.1093/infdis/jiw159.
15. Bradshaw, C.S., and Brotman, R.M. (2015). Making inroads into improving treatment of bacterial vaginosis - striving for long-term cure. BMC Infect. Dis. *15*, 292. https://doi.org/10.1186/s12879-015-1027-4.
16. Schwebke, J.R., Lensing, S.Y., Lee, J., Muzny, C.A., Pontius, A., Woznicki, N., Aguin, T., and Sobel, J.D. (2021). Treatment of Male Sexual Partners of Women With Bacterial Vaginosis: A Randomized, Double-Blind, Placebo-Controlled Trial. Clin. Infect. Dis. *73*, e672–e679. https://doi.org/10.1093/cid/ciaa1903.
17. Bradshaw, C.S., Morton, A.N., Hocking, J., Garland, S.M., Morris, M.B., Moss, L.M., Horvath, L.B., Kuzevska, I., and Fairley, C.K. (2006). High recurrence rates of bacterial vaginosis over the course of 12 months after oral metronidazole therapy and factors associated with recurrence. J. Infect. Dis. *193*, 1478–1486. https://doi.org/10.1086/503780.
18. Hemmerling, A., Wierzbicki, M.R., Armstrong, E., and Cohen, C.R. (2024). Response to Antibiotic Treatment of Bacterial Vaginosis Predicts the Effectiveness of LACTIN-V (Lactobacillus crispatus CTV-05) in the Prevention of Recurrent Disease. Sex. Transm. Dis. *51*, 437–440. https://doi.org/10.1097/OLQ.0000000000001962.
19. Munoz, A., Hayward, M.R., Bloom, S.M., Rocafort, M., Ngcapu, S., Mafunda, N.A., Xu, J., Xulu, N., Dong, M., Dong, K.L., et al. (2021). Modeling the temporal dynamics of cervicovaginal microbiota identifies targets that may promote reproductive health. Microbiome *9*, 163. https://doi.org/10.1186/s40168-021-01096-9.
20. Tamarelle, J., Shardell, M.D., Ravel, J., and Brotman, R.M. (2022). Factors Associated With Incidence and Spontaneous Clearance of Molecular-Bacterial Vaginosis: Results From a Longitudinal Frequent-Sampling Observational Study. Sex. Transm. Dis. *49*, 649–656. https://doi.org/10.1097/OLQ.0000000000001662.
21. DiGiulio, D.B., Callahan, B.J., McMurdie, P.J., Costello, E.K., Lyell, D.J., Robaczewska, A., Sun, C.L., Goltsman, D.S.A., Wong, R.J., Shaw, G., et al. (2015). Temporal and spatial variation of the human microbiota during pregnancy. Proc. Natl. Acad. Sci. USA *112*, 11060–11065. https://doi.org/10.1073/pnas.1502875112.
22. van Houdt, R., Ma, B., Bruisten, S.M., Speksnijder, A.G.C.L., Ravel, J., and de Vries, H.J.C. (2018). Lactobacillus iners-dominated vaginal microbiota is associated with increased susceptibility to Chlamydia trachomatis infection in Dutch women: a case-control study. Sex. Transm. Infect. *94*, 117–123. https://doi.org/10.1136/sextrans-2017-053133.
23. Colbert, L.E., El Alam, M.B., Wang, R., Karpinets, T., Lo, D., Lynn, E.J., Harris, T.A., Elnaggar, J.H., Yoshida-Court, K., Tomasic, K., et al. (2023). Tumor-resident Lactobacillus iners confer chemoradiation resistance through lactate-induced metabolic rewiring. Cancer Cell *41*, 1945–1962.e11. https://doi.org/10.1016/j.ccell.2023.09.012.

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

24. Norenhag, J., Du, J., Olovsson, M., Verstraelen, H., Engstrand, L., and Brusselaers, N. (2020). The vaginal microbiota, human papillomavirus and cervical dysplasia: a systematic review and network meta-analysis. BJOG *127*, 171–180. https://doi.org/10.1111/1471-0528.15854.

25. Kindinger, L.M., Bennett, P.R., Lee, Y.S., Marchesi, J.R., Smith, A., Cacciatore, S., Holmes, E., Nicholson, J.K., Teoh, T.G., and MacIntyre, D.A. (2017). The interaction between vaginal microbiota, cervical length, and vaginal progesterone treatment for preterm birth risk. Microbiome *5*, 6. https://doi.org/10.1186/s40168-016-0223-9.

26. Srinivasan, S., Liu, C., Mitchell, C.M., Fiedler, T.L., Thomas, K.K., Agnew, K.J., Marrazzo, J.M., and Fredricks, D.N. (2010). Temporal variability of human vaginal bacteria and relationship with bacterial vaginosis. PLoS One *5*, e10197. https://doi.org/10.1371/journal.pone.0010197.

27. Ravel, J., Brotman, R.M., Gajer, P., Ma, B., Nandy, M., Fadrosh, D.W., Sakamoto, J., Koenig, S.S., Fu, L., Zhou, X., et al. (2013). Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. Microbiome *1*, 29. https://doi.org/10.1186/2049-2618-1-29.

28. Verwijs, M.C., Agaba, S.K., Darby, A.C., and Van De Wijgert, J.H.H.M. (2020). Impact of oral metronidazole treatment on the vaginal microbiota and correlates of treatment failure. Am. J. Obstet. Gynecol. *222*, 157.e1–157.e13. https://doi.org/10.1016/j.ajog.2019.08.008.

29. Joag, V., Obila, O., Gajer, P., Scott, M.C., Dizzell, S., Humphrys, M., Shahabi, K., Huibner, S., Shannon, B., Tharao, W., et al. (2019). Impact of Standard Bacterial Vaginosis Treatment on the Genital Microbiota, Immune Milieu, and Ex Vivo Human Immunodeficiency Virus Susceptibility. Clin. Infect. Dis. *68*, 1675–1683. https://doi.org/10.1093/cid/ciy762.

30. Mitchell, C., Manhart, L.E., Thomas, K., Fiedler, T., Fredricks, D.N., and Marrazzo, J. (2012). Behavioral predictors of colonization with Lactobacillus crispatus or Lactobacillus jensenii after treatment for bacterial vaginosis: a cohort study. Infect. Dis. Obstet. Gynecol. *2012*, 706540. https://doi.org/10.1155/2012/706540.

31. Ferris, M.J., Norori, J., Zozaya-Hinchliffe, M., and Martin, D.H. (2007). Cultivation-independent analysis of changes in bacterial vaginosis flora following metronidazole treatment. J. Clin. Microbiol. *45*, 1016–1018. https://doi.org/10.1128/JCM.02085-06.

32. Zhu, M., Frank, M.W., Radka, C.D., Jeanfavre, S., Xu, J., Tse, M.W., Pacheco, J.A., Kim, J.S., Pierce, K., Deik, A., et al. (2024). Vaginal Lactobacillus fatty acid response mechanisms reveal a metabolite-targeted strategy for bacterial vaginosis treatment. Cell *187*, 5413–5430.e29. https://doi.org/10.1016/j.cell.2024.07.029.

33. Nilsen, T., Swedek, I., Lagenaur, L.A., and Parks, T.P. (2020). Novel Selective Inhibition of Lactobacillus iners by Lactobacillus-Derived Bacteriocins. Appl. Environ. Microbiol. *86*. e01594-20. https://doi.org/10.1128/AEM.01594-20.

34. Lev-Sagie, A., Goldman-Wohl, D., Cohen, Y., Dori-Bachash, M., Leshem, A., Mor, U., Strahilevitz, J., Moses, A.E., Shapiro, H., Yagel, S., et al. (2019). Vaginal microbiome transplantation in women with intractable bacterial vaginosis. Nat. Med. *25*, 1500–1504. https://doi.org/10.1038/s41591-019-0600-6.

35. Bloom, S.M., Mafunda, N.A., Woolston, B.M., Hayward, M.R., Frempong, J.F., Abai, A.B., Xu, J., Mitchell, A.J., Westergaard, X., Hussain, F.A., et al. (2022). Cysteine dependence of Lactobacillus iners is a potential therapeutic target for vaginal microbiota modulation. Nat. Microbiol. *7*, 434–450. https://doi.org/10.1038/s41564-022-01070-7.

36. Yockey, L.J., Hussain, F.A., Bergerat, A., Reissis, A., Worrall, D., Xu, J., Gomez, I., Bloom, S.M., Mafunda, N.A., Kelly, J., et al. (2022). Screening and characterization of vaginal fluid donations for vaginal microbiota transplantation. Sci. Rep. *12*, 17948. https://doi.org/10.1038/s41598-022-22873-y.

37. Cohen, C.R., Wierzbicki, M.R., French, A.L., Morris, S., Newmann, S., Reno, H., Green, L., Miller, S., Powell, J., Parks, T., et al. (2020). Randomized Trial of Lactin-V to Prevent Recurrence of Bacterial Vaginosis. N. Engl. J. Med. *382*, 1906–1915. https://doi.org/10.1056/NEJMoa1915254.

38. Ravel, J., Simmons, S., Jaswa, E.G., Gottfried, S., Greene, M., Kellogg-Spadt, S., Gevers, D., and Harper, D.M. (2025). Impact of a multi-strain L. crispatus-based vaginal synbiotic on the vaginal microbiome: a randomized placebo-controlled trial. NPJ Biofilms Microbiomes *11*, 158. https://doi.org/10.1038/s41522-025-00788-6.

39. Chetty, C., Mafunda, N.A., Happel, A.-U., Khan, A., Demidkina, B.C., Yende-Zuma, N., Saidi, Y., Polliah, A.M., Lewis, L., Osman, F., et al. (2025). Randomized Trial of Multi-Strain Lactobacillus crispatus Vaginal Live Biotherapeutic Products after Antibiotic Therapy for Bacterial Vaginosis: Study Protocol for VIBRANT (Vaginal Live Biotherapeutic RANdomized Trial). Contemp. Clin. Trials Commun. *48*, 101554.

40. Hemmerling, A., Harrison, W., Schroeder, A., Park, J., Korn, A., Shiboski, S., and Cohen, C.R. (2009). Phase 1 dose-ranging safety trial of Lactobacillus crispatus CTV-05 for the prevention of bacterial vaginosis. Sex. Transm. Dis. *36*, 564–569. https://doi.org/10.1097/OLQ.0b013e3181a74924.

41. Antonio, M.A., Hawes, S.E., and Hillier, S.L. (1999). The identification of vaginal Lactobacillus species and the demographic and microbiologic characteristics of women colonized by these species. J. Infect. Dis. *180*, 1950–1956. https://doi.org/10.1086/315109.

42. Antonio, M.A.D., and Hillier, S.L. (2003). DNA fingerprinting of Lactobacillus crispatus strain CTV-05 by repetitive element sequence-based PCR analysis in a pilot study of vaginal colonization. J. Clin. Microbiol. *41*, 1881–1887. https://doi.org/10.1128/JCM.41.5.1881-1887.2003.

43. Lagenaur, L.A., Hemmerling, A., Chiu, C., Miller, S., Lee, P.P., Cohen, C.R., and Parks, T.P. (2021). Connecting the Dots: Translating the Vaginal Microbiome Into a Drug. J. Infect. Dis. *223*, S296–S306. https://doi.org/10.1093/infdis/jiaa676.

44. Amsel, R., Totten, P.A., Spiegel, C.A., Chen, K.C.S., Eschenbach, D., and Holmes, K.K. (1983). Nonspecific vaginitis: Diagnostic criteria and microbial and epidemiologic associations. Am. J. Med. *74*, 14–22. https://doi.org/10.1016/0002-9343(83)91112-9.

45. Nugent, R.P., Krohn, M.A., and Hillier, S.L. (1991). Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. J. Clin. Microbiol. *29*, 297–301. https://doi.org/10.1128/jcm.29.2.297-301.1991.

46. Bradshaw, C.S., Plummer, E.L., Muzny, C.A., Mitchell, C.M., Fredricks, D.N., Herbst-Kralovetz, M.M., and Vodstrcil, L.A. (2025). Bacterial vaginosis. Nat. Rev. Dis. Primers *11*, 43. https://doi.org/10.1038/s41572-025-00626-1.

47. Robert, P., and Escoufier, Y. (1976). A Unifying Tool for Linear Multivariate Statistical Methods: The RV- Coefficient. J. R. Stat. Soc. C *25*, 257–265. https://doi.org/10.2307/2347233.

48. Josse, J., and Holmes, S. (2016). Measuring multivariate association and beyond. Stat. Surv. *10*, 132–167. https://doi.org/10.1214/16-SS116.

49. Sankaran, K., and Holmes, S.P. (2019). Latent variable modeling for the microbiome. Biostatistics *20*, 599–614. https://doi.org/10.1093/biostatistics/kxy018.

50. Symul, L., Jeganathan, P., Costello, E.K., France, M., Bloom, S.M., Kwon, D.S., Ravel, J., Relman, D.A., and Holmes, S. (2023). Sub-communities of the vaginal microbiota in pregnant and non-pregnant women. Proc. Biol. Sci. *290*, 20231461. https://doi.org/10.1098/rspb.2023.1461.

51. Ravel, J., Gajer, P., Abdo, Z., Schneider, G.M., Koenig, S.S.K., McCulle, S.L., Karlebach, S., Gorle, R., Russell, J., Tacket, C.O., et al. (2011). Vaginal microbiome of reproductive-age women. Proc. Natl. Acad. Sci. USA *108*, 4680–4687. https://doi.org/10.1073/pnas.1002611107.

52. France, M.T., Ma, B., Gajer, P., Brown, S., Humphrys, M.S., Holm, J.B., Waetjen, L.E., Brotman, R.M., and Ravel, J. (2020). VALENCIA: a nearest centroid classification method for vaginal microbial communities based on composition. Microbiome *8*, 166. https://doi.org/10.1186/s40168-020-00934-6.

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

53. Blei, D.M., Ng, A.Y., and Jordan, M.I. (2003). Latent Dirichlet Allocation. J. Mach. Learn. Res. *3*, 993–1022.

54. Smith, B.J., Li, X., Shi, Z.J., Abate, A., and Pollard, K.S. (2022). Scalable Microbial Strain Inference in Metagenomic Data Using StrainFacts. Front. Bioinform. *2*, 867386. https://doi.org/10.3389/fbinf.2022.867386.

55. Shih, J., Bloom, S.M., Xu, J., Mitchell, C.M., Elsherbini, J., and Kwon, D.S. (2025). StrainFacts accurately quantifies both endogenous and live biotherapeutic product strain abundances in simulated and clinical vaginal microbiota samples. Preprint at bioRxiv. https://doi.org/10.1101/2025.08.15.670563.

56. Jolicoeur, P., and Mosimann, J.E. (1960). Size and Shape Variation in The Painted Turtle. A Principal Component Analysis. Growth *24*, 339–354.

57. Masson, L., Mlisana, K., Little, F., Werner, L., Mkhize, N.N., Ronacher, K., Gamieldien, H., Williamson, C., Mckinnon, L.R., Walzl, G., et al. (2014). Defining genital tract cytokine signatures of sexually transmitted infections and bacterial vaginosis in women at high risk of HIV infection: a cross-sectional study. Sex. Transm. Infect. *90*, 580–587. https://doi.org/10.1136/sextrans-2014-051601.

58. Masson, L., Barnabas, S., Deese, J., Lennard, K., Dabee, S., Gamieldien, H., Jaumdally, S.Z., Williamson, A.-L., Little, F., Van Damme, L., et al. (2019). Inflammatory cytokine biomarkers of asymptomatic sexually transmitted infections and vaginal dysbiosis: a multicentre validation study. Sex. Transm. Infect. *95*, 5–12. https://doi.org/10.1136/sextrans-2017-053506.

59. Abdi, H., O'Toole, A.J., Valentin, D., and Edelman, B. (2005). DISTATIS: The Analysis of Multiple Distance Matrices. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops (IEEE), p. 42. https://doi.org/10.1109/CVPR.2005.445.

60. Abdi, H., Dunlop, J.P., and Williams, L.J. (2009). How to compute reliability estimates and display confidence and tolerance intervals for pattern classifiers using the Bootstrap and 3-way multidimensional scaling (DISTATIS). Neuroimage *45*, 89–95. https://doi.org/10.1016/j.neuroimage.2008.11.008.

61. Bougeard, S., Qannari, E.M., and Rose, N. (2011). Multiblock redundancy analysis: interpretation tools and application in epidemiology. J. Chemom. *25*, 467–475. https://doi.org/10.1002/cem.1392.

62. Brandolini-Bunlon, M., Pétéra, M., Gaudreau, P., Comte, B., Bougeard, S., and Pujos-Guillot, E. (2019). Multi-block PLS discriminant analysis for the joint analysis of metabolomic and epidemiological data. Metabolomics *15*, 134. https://doi.org/10.1007/s11306-019-1598-y.

63. Farmer, N.M., Benezra, A., Maki, K.A., Ishaq, S.L., and Kozik, A.J. (2025). Prioritizing precision: guidelines for the better use of population descriptors in human microbiome research. mSystems *10*, e0064025. https://doi.org/10.1128/msystems.00640-25.

64. National Academies of Sciences, Engineering, and Medicine (2023). Using Population Descriptors in Genetics and Genomics Research: A New Framework for an Evolving Field (National Academies Press). https://doi.org/10.17226/26902.

65. Armstrong, E., Hemmerling, A., Miller, S., Burke, K.E., Newmann, S.J., Morris, S.R., Reno, H., Huibner, S., Kulikova, M., Nagelkerke, N., et al. (2022). Sustained effect of LACTIN-V (Lactobacillus crispatus CTV-05) on genital immunology following standard bacterial vaginosis treatment: results from a randomised, placebo-controlled trial. Lancet Microbe *3*, e435–e442. https://doi.org/10.1016/S2666-5247(22)00043-X.

66. Aldridge, K.E., Ashcraft, D., Cambre, K., Pierson, C.L., Jenkins, S.G., and Rosenblatt, J.E. (2001). Multicenter survey of the changing in vitro antimicrobial susceptibilities of clinical isolates of Bacteroides fragilis group, Prevotella, Fusobacterium, Porphyromonas, and Peptostreptococcus species. Antimicrob. Agents Chemother. *45*, 1238–1243. https://doi.org/10.1128/AAC.45.4.1238-1243.2001.

67. Alauzet, C., Mory, F., Teyssier, C., Hallage, H., Carlier, J.P., Grollier, G., and Lozniewski, A. (2010). Metronidazole resistance in Prevotella spp. and description of a new nim gene in Prevotella baroniae. Antimicrob. Agents Chemother. *54*, 60–64. https://doi.org/10.1128/AAC.01003-09.

68. Mitchell, C.M., Hitti, J.E., Agnew, K.J., and Fredricks, D.N. (2009). Comparison of oral and vaginal metronidazole for treatment of bacterial vaginosis in pregnancy: impact on fastidious bacteria. BMC Infect. Dis. *9*, 89. https://doi.org/10.1186/1471-2334-9-89.

69. Holm, J.B., France, M.T., Ma, B., McComb, E., Robinson, C.K., Mehta, A., Tallon, L.J., Brotman, R.M., and Ravel, J. (2020). Comparative Metagenome-Assembled Genome Analysis of "Candidatus Lachnocurva vaginae", Formerly Known as Bacterial Vaginosis-Associated Bacterium-1 (BVAB1). Front. Cell. Infect. Microbiol. *10*, 117. https://doi.org/10.3389/fcimb.2020.00117.

70. Fredricks, D.N., Fiedler, T.L., and Marrazzo, J.M. (2005). Molecular Identification of Bacteria Associated with Bacterial Vaginosis. N. Engl. J. Med. *353*, 1899–1911. https://doi.org/10.1056/nejmoa043802.

71. Amabebe, E., and Anumba, D.O.C. (2018). Psychosocial Stress, Cortisol Levels, and Maintenance of Vaginal Health. Front. Endocrinol. (Lausanne) *9*, 568. https://doi.org/10.3389/fendo.2018.00568.

72. McCosh, R.B., O'Bryne, K.T., Karsch, F.J., and Breen, K.M. (2022). Regulation of the gonadotropin-releasing hormone neuron during stress. J. Neuroendocrinol. *34*, e13098. https://doi.org/10.1111/jne.13098.

73. Keshavarz, H., Duffy, S.W., Sadeghi-Hassanabadi, A., Zolghadr, Z., and Oboodi, B. (2001). Risk factors for and relationship between bacterial vaginosis and cervicitis in a high risk population for cervicitis in Southern Iran. Eur. J. Epidemiol. *17*, 89–95. https://doi.org/10.1023/A:1010935723248.

74. Wenman, W.M., Joffres, M.R., and Tataryn, I.V.; Edmonton Perinatal Infections Group (2004). A prospective cohort study of pregnancy risk factors and birth outcomes in Aboriginal women. CMAJ *171*, 585–589. https://doi.org/10.1503/cmaj.1031730.

75. Dai, Q., Hu, L., Jiang, Y., Shi, H., Liu, J., Zhou, W., Shen, C., and Yang, H. (2010). An epidemiological survey of bacterial vaginosis, vulvovaginal candidiasis and trichomoniasis in the Tibetan area of Sichuan Province, China. Eur. J. Obstet. Gynecol. Reprod. Biol. *150*, 207–209. https://doi.org/10.1016/j.ejogrb.2010.02.027.

76. Virtanen, S., Rantsi, T., Virtanen, A., Kervinen, K., Nieminen, P., Kalliala, I., and Salonen, A. (2019). Vaginal Microbiota Composition Correlates Between Pap Smear Microscopy and Next Generation Sequencing and Associates to Socioeconomic Status. Sci. Rep. *9*, 7750. https://doi.org/10.1038/s41598-019-44157-8.

77. Liu, Z., Yu, W., Sun, T., Li, M., Li, X., Qin, L., Liu, X., Bian, Y., Zhao, S., Zhao, Q., et al. (2025). Ambient temperature affects the composition of the vaginal microbiome, and temperature-sensitive vaginal microbes influence assisted reproductive technology outcomes. Microbiome *14*, 3. https://doi.org/10.1186/s40168-025-02212-9.

78. Vodstrcil, L.A., Plummer, E.L., Fairley, C.K., Hocking, J.S., Law, M.G., Petoumenos, K., Bateson, D., Murray, G.L., Donovan, B., Chow, E.P.F., et al. (2025). Male-Partner Treatment to Prevent Recurrence of Bacterial Vaginosis. N. Engl. J. Med. *392*, 947–957. https://doi.org/10.1056/NEJMoa2405404.

79. Zenilman, J.M., Weisman, C.S., Rompalo, A.M., Ellish, N., Upchurch, D.M., Hook, E.W.I., and Celentano, D. (1995). Condom Use to Prevent Incident STDs: The Validity of Self-Reported Condom Use. Sex. Transm. Dis. *22*, 15–21. https://doi.org/10.1097/00007435-199501000-00003.

80. Schroder, K.E.E., Carey, M.P., and Vanable, P.A. (2003). Methodological challenges in research on sexual risk behavior: I. Item content, scaling, and data analytical options. Ann. sBehav. Med. *26*, 76–103. https://doi.org/10.1207/S15324796ABM2602_02.

81. Schroder, K.E.E., Carey, M.P., and Vanable, P.A. (2003). Methodological challenges in research on sexual risk behavior: II. Accuracy of self-reports. Ann. Behav. Med. *26*, 104–123. https://doi.org/10.1207/S15324796ABM2602_03.

82. Jewanraj, J., Ngcapu, S., Osman, F., Mtshali, A., Singh, R., Mansoor, L.E., Abdool Karim, S.S., Abdool Karim, Q., Passmore, J.S., and Liebenberg, L.J.P. (2020). The Impact of Semen Exposure on the Immune and Microbial Environments of the Female Genital Tract. Front. Reprod. Health *2*, 566559. https://doi.org/10.3389/frph.2020.566559.

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

83. Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N., and Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc. Natl. Acad. Sci. USA *108*, 4516–4522. https://doi.org/10.1073/pnas.1000080107.

84. Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods *7*, 335–336. https://doi.org/10.1038/nmeth.f.303.

85. Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods *13*, 581–583. https://doi.org/10.1038/nmeth.3869.

86. McMurdie, P.J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One *8*, e61217. https://doi.org/10.1371/journal.pone.0061217.

87. Aragon, T.J. (2004). epitools: Epidemiology Tools. https://doi.org/10.32614/CRAN.package.epitools.

88. Shi, Z.J., Dimitrov, B., Zhao, C., Nayfach, S., and Pollard, K.S. (2022). Fast and accurate metagenotyping of the human gut microbiome with GT-Pro. Nat. Biotechnol. *40*, 507–516. https://doi.org/10.1038/s41587-021-01102-3.

89. Shi, Z.J., Nayfach, S., and Pollard, K.S. (2023). Maast: genotyping thousands of microbial strains efficiently. Genome Biol. *24*, 186. https://doi.org/10.1186/s13059-023-03030-8.

90. Schwengers, O., Jelonek, L., Dieckmann, M.A., Beyvers, S., Blom, J., and Goesmann, A. (2021). Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microb. Genomics *7*, 000685. https://doi.org/10.1099/mgen.0.000685.

91. Wick, R.R., Judd, L.M., Cerdeira, L.T., Hawkey, J., Méric, G., Vezina, B., Wyres, K.L., and Holt, K.E. (2021). Trycycler: consensus long-read assemblies for bacterial genomes. Genome Biol. *22*, 266. https://doi.org/10.1186/s13059-021-02483-z.

92. Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner. Report Number: LBNL-7065E. https://www.osti.gov/biblio/1241166.

93. Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P.A. (2019). Assembly of long, error-prone reads using repeat graphs. Nat. Biotechnol. *37*, 540–546. https://doi.org/10.1038/s41587-019-0072-8.

94. Wick, R.R., and Holt, K.E. (2019). Benchmarking of long-read assemblers for prokaryote whole genome sequencing. F1000Res *8*, 2138. https://doi.org/10.12688/f1000research.21782.4.

95. Vaser, R., and Šikić, M. (2021). Time- and memory-efficient genome assembly with Raven. Nat Comput. Sci. *1*, 332–336. https://doi.org/10.1038/s43588-021-00073-4.

96. Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., and Phillippy, A.M. (2017). Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Res. *27*, 722–736. https://doi.org/10.1101/gr.215087.116.

97. Wick, R.R., and Holt, K.E. (2022). Polypolish: Short-read polishing of long-read bacterial genome assemblies. PLoS Comput. Biol. *18*, e1009802. https://doi.org/10.1371/journal.pcbi.1009802.

98. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. *17*, 10. https://doi.org/10.14806/ej.17.1.200.

99. Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. *37*, 907–915. https://doi.org/10.1038/s41587-019-0201-4.

100. Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P., and Parks, D.H. (2022). GTDB-Tk v2: memory friendly classification with the genome taxonomy database. Bioinformatics *38*, 5315–5316. https://doi.org/10.1093/bioinformatics/btac672.

101. Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics *11*, 119. https://doi.org/10.1186/1471-2105-11-119.

102. Eddy, S.R. (2023). HMMER. Version 3.4. http://eddylab.org/software/hmmer/Userguide.pdf.

103. Katoh, K., Rozewicki, J., and Yamada, K.D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief. Bioinform. *20*, 1160–1166. https://doi.org/10.1093/bib/bbx108.

104. Kuraku, S., Zmasek, C.M., Nishimura, O., and Katoh, K. (2013). aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. Nucleic Acids Res. *41*, W22–W28. https://doi.org/10.1093/nar/gkt389.

105. Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., and Stamatakis, A. (2019). RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics *35*, 4453–4455. https://doi.org/10.1093/bioinformatics/btz305.

106. Morgan, M., Obenchain, V., Hester, J., and Pagès, H. (2017). SummarizedExperiment. Bioconductor. https://doi.org/10.18129/B9.BIOC.SUMMARIZEDEXPERIMENT.

107. Ramos, M., Morgan, M., Shepherd, L., Pagès, H., Carey, V.J., Waldron, L., and MultiAssay SIG. (2017). MultiAssayExperiment. Bioconductor. https://doi.org/10.18129/B9.BIOC.MULTIASSAYEXPERIMENT.

108. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the Tidyverse. J. Open Source Softw. *4*, 1686. https://doi.org/10.21105/joss.01686.

109. Firke, S. (2024). janitor: Simple Tools for Examining and Cleaning Dirty Data. Version 2.2.1.9000. https://cran.r-project.org/web/packages/janitor/index.html.

110. Iannone, R., Cheng, J., Schloerke, B., Haughton, S., Hughes, E., Lauer, A., François, R., Seo, J., Brevoort, K., and Roy, O. (2025). gt: Easily Create Presentation-Ready Display Tables. R package version 1.2.0.9000. https://cran.r-project.org/web/packages/gt/index.html.

111. Fukuyama, J., Sankaran, K., and Symul, L. (2024). alto: Aligns Topics across LDA Models. Version 0.1.0. https://lasy.github.io/alto/.

112. Grün, B., and Hornik, K. (2011). **topicmodels**: An R Package for Fitting Topic Models. J. Stat. Soft. *40*, 1–30. https://doi.org/10.18637/jss.v040.i13.

113. Sjoberg, D., Whiting, K., Curry, M., Lavery, J., and Larmarange, J. (2021). Reproducible Summary Tables with the gtsummary Package. R J. *13*, 570. https://doi.org/10.32614/RJ-2021-053.

114. Dray, S., and Dufour, A.-B. (2007). The **ade4** Package: Implementing the Duality Diagram for Ecologists. J. Stat. Softw. *22*. https://doi.org/10.18637/jss.v022.i04.

115. Kassambara, A., and Mundt, F. (2016). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. https://doi.org/10.32614/CRAN.package.factoextra.

116. Beaton, D., Yu, J.-C., Guillemot, V., and Abdi, H. (2013). DistatisR: DiSTATIS Three Way Metric Multidimensional Scaling. https://doi.org/10.32614/CRAN.package.DistatisR.

117. Scherer, R. (2010). PropCIs: Various Confidence Interval Methods for Proportions. https://doi.org/10.32614/CRAN.package.PropCIs.

118. Liland, K.H., Mevik, B.-H., and Wehrens, R. (2025). pls: Partial Least Squares and Principal Component Regression. R package version 2.8-5.1. https://cran.r-project.org/web/packages/pls/pls.pdf.

119. Brandolini-Bunlon, M., Bougeard, S., Petera, M., and Pujos-Guillot, E. (2019). packMBPLSDA: Multi-Block Partial Least Squares Discriminant Analysis. https://doi.org/10.32614/CRAN.package.packMBPLSDA.

120. Joshi, N.A., and Fass, J.N. (2011). Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files. Version 1.33. GitHub. https://github.com/najoshi/sickle.

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

121. Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Version 0.12.1. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

122. Wick, R.R., Judd, L.M., Gorrie, C.L., and Holt, K.E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput. Biol. *13*, e1005595. https://doi.org/10.1371/journal.pcbi.1005595.

123. Schwengers, O. (2024). Bakta database. Version 5.1. Zenodo. https://doi.org/10.5281/ZENODO.10522951.

124. Manni, M., Berkeley, M.R., Seppey, M., Simão, F.A., and Zdobnov, E.M. (2021). BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. Mol. Biol. Evol. *38*, 4647–4654. https://doi.org/10.1093/molbev/msab199.

125. Anahtar, M.N., Bowman, B.A., and Kwon, D.S. (2016). Efficient Nucleic Acid Extraction and 16S rRNA Gene Sequencing for Bacterial Community Characterization. J. Vis. Exp. *110*, 53939. https://doi.org/10.3791/53939.

126. Baym, M., Kryazhimskiy, S., Lieberman, T.D., Chung, H., Desai, M.M., and Kishony, R. (2015). Inexpensive multiplexed library preparation for megabase-sized genomes. PLoS One *10*, e0128036. https://doi.org/10.1371/journal.pone.0128036.

127. Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Solymos, P., Stevens, M.H.H., Szoecs, E., et al. (2001). vegan: Community Ecology Package. https://doi.org/10.32614/CRAN.package.vegan.

128. Fukuyama, J., Sankaran, K., and Symul, L. (2023). Multiscale analysis of count data through topic alignment. Biostatistics *24*, 1045–1065. https://doi.org/10.1093/biostatistics/kxac018.

129. Bougeard, S., and Dray, S. (2018). Supervised Multiblock Analysis in *R* with the **ade4** Package. J. Stat. Softw. *86*. https://doi.org/10.18637/jss.v086.i01.

130. Abdi, H., Williams, L.J., Valentin, D., and Bennani-Dosse, M. (2012). STATIS and DISTATIS: optimum multitable principal component analysis and three way metric multidimensional scaling. WIREs Computational Stats. *4*, 124–167. https://doi.org/10.1002/wics.198.

131. Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc., B (Methodol.) *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Lactobacillus crispatus strain CTV-05	Osel, Inc.	CTV-05
Vaginal bacterial primary isolates	This study	See Data S3
Biological samples		
Human vaginal swab samples	Cohen et al.[37]	N/A
Chemicals, peptides, and recombinant proteins		
Lactobacillus MRS agar	Hardy Diagnostics	#G117
Columbia Blood Agar ("CBA")	Hardy Diagnostics	#A16
CDC Anaerobe Laked Sheep Blood Agar with Kanamycin and Vancomycin ("LKV")	BD BBL™ Prepared Plated Media	#221846
Lactobacillus MRS broth	BD	#288130
Wilkins-Chalgren Anaerobe Broth	Thermo Scientific™ Oxoid™	#CM0643B
HEPES	Fisher Scientific	#BP310-500
Proteose peptone no. 3	BD Biosciences	#BD 211693
Dextrose	Fisher Chemical™	#D16-500
Yeast extract solution	Gibco	#18180-059
Heat-inactivated horse serum	Gibco	#26050070
Dulbecco's Phosphate Buffered Saline ("PBS")	Millipore Sigma	#D8537
Phenol:Chloroform:IAA, 25:24:1, pH 6.6	Invitrogen	#AM9730 (since discontinued)
Sodium Dodecyl Sulfate 20% Solution	Fisher Scientific	#BP1311-200
EDTA	Invitrogen	#AM9260G
2-Propanol	Sigma	# I9516-500ML
Sodium Acetate, pH 5.5	Life Technologies	#AM9740
TE Buffer	Promega	#V6321
QIAamp 96 DNA QIAcube HT kit	Qiagen	#51331
Q5 reaction buffer	NEB	#B9027
Ultrapure dNTP mix	NEB	#N0447
Q5 high-fidelity DNA polymerase	NEB	#M0491
Invitrogen UltraPure Water – 0.1-μm Membrane-Filtered Water	Invitrogen	#10977
SYBR™ Green I Nucleic Acid Gel Stain, 10,000X concentrate in DMSO	Invitrogen	#S7653
HighPrep PCR magnetic beads	MagBio Genomics	#AC-60050
5M Lithium Chloride	Molecular Dimensions	#MD2-100-43
Multiscreen® 96 well Plate, hydrophilic PVDF membrane (pore size 0.22 μm, sterile)	EMD Millipore	#MSGVS2210
Critical commercial assays		
Starplex™ Scientific Multitrans™ Collection and Transportation System	Starplex™ Scientific	S1600
ESwab® Liquid Based Collection and Transport System	Copan	ESwab 480C®
QIAquick PCR purification kit	Qiagen	#28104
MiSeq Reagent Kit v2, 300 Cycles (PE)	Illumina	#MS-102-2002
Nextera DNA Library Preparation Kit	Illumina	#20034211
KAPA HiFi Library Amplification Kit	Kapa Biosystems	#KK2602
MasterPure Gram Positive DNA Purification Kit	Biosearch Technologies	#MGP04100

(*Continued on next page*)

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Oxford Nanopore Sequencing Rapid Barcoding Kit	Oxford Nanopore Technologies	#SQK-RBK004 (since discontinued)
MinION Flow Cell using R9.4.1 chemistry	Oxford Nanopore Technologies	#FLO-MIN106D (since discontinued)
MinION Sequencing Device	Oxford Nanopore Technologies	#MIN-101B
Custom 20-plex High Sensitivity Luminex Assay Kit	EMD Millipore	Gosmann et al.[6]
Deposited data		
Metagenomic sequence data, bacterial 16S rRNA gene sequence data, isolate shotgun genomic read sequence data, and isolate genome assemblies.	NCBI	BioProject: PRJNA1303956
CTV-05 closed genome and the Illumina short-read shotgun genomic sequences and Oxford Nanopore long-read shotgun genomic sequences from which it was assembled	NCBI	BioSample: SAMN55039759 (part of BioProject: PRJNA1303956, above)
Clinical data and processed and transformed omics data	Zenodo	Zenodo: https://doi.org/10.5281/zenodo.17755173
Analysis code	Zenodo	Zenodo: https://doi.org/10.5281/zenodo.18524760
Oligonucleotides		
Bacterial 16S rRNA gene 515F primer AATGATACGG CGACCACCGAGACGTACGTACGGTGTGCCAGCM GCCGCGGTAA (underlined sequence represents the region of complementarity to the bacterial gene)	Caporaso et al.[83]; primer manufactured by IDT	N/A
Bacterial 16S rRNA gene 806R primer CAAGCAGAA GACGGCATACGAGATXXXXXXXXXXXXAGTCAGTC AGCCGGACTACHVGGGTWTCTAAT (underlined sequence represents the region of complementarity to the bacterial gene; barcode positions in the primer are indicated by X)	Caporaso et al.[83]; primer manufactured by IDT	N/A
Custom Earth Microbiome Project Read 1 sequencing primer: ACGTACGTACGGTGTGCCAGCMGCCGCGGTAA	Caporaso et al.[83]; primer manufactured by IDT	N/A
Custom Earth Microbiome Project Read 2 sequencing primer: ACGTACGTACCCGGACTACHVGGGTWTCTAAT	Caporaso et al.[83]; primer manufactured by IDT	N/A
Custom Earth Microbiome Project index sequencing primer: ATTAGAWACCCBDGTAGTCCGGCTGACTGACT	Caporaso et al.[83]; primer manufactured by IDT	N/A
Total bacterial 16S qPCR primer, forward: AGAGTTTG ATCCTGGCTCAG	Cohen et al.[37]	N/A
Total bacterial 16S qPCR primer, reverse: GCTGCCT CCCGTAGGAGT	Cohen et al.[37]	N/A
Software and algorithms		
QIIME 1 version 1.9.188	Caporaso et al.[84]	N/A
dada2 version 1.6.0 (R package)	Callahan et al.[85]	N/A
Ribosomal Database Project (RDP) training database rdp_train_set_16.fa.gz	https://www.mothur.org/wiki/RDP_reference_files	N/A
phyloseq (R package)	McMurdie and Holmes[86]	N/A
vegan (R package)	https://vegandevs.github.io/vegan/	N/A
epitools (R package)	Aragon[87]	N/A
GT-Pro v1.0.1	Shi et al.[88]	N/A
MAAST	Shi et al.[89]	N/A
StrainFacts	Smith et al.[54]	N/A
Bakta v1.9.4	Schwengers et al.[90]	N/A
MinKNOW v2.2	Oxford Nanopore Technologies	N/A
Luminex xPONENT software	Diasorin	N/A
Filtlong v0.2.1	https://github.com/rrwick/Filtlong	N/A
Trycycler v0.5.3	Wick et al.[91]	N/A

(*Continued on next page*)

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BBMap v39.00	Bushnell[92]	N/A
Flye v2.9.1	Kolmogorov et al.[93]	N/A
Minipolish v0.1.3	Wick and Holt[94]	N/A
Raven v1.8.1	Vaser and Šikić[95]	N/A
Canu v2.2	Koren et al.[96]	N/A
Medaka v1.7.2	https://github.com/nanoporetech/medaka	N/A
Polypolish v0.5.0	Wick and Holt[97]	N/A
Cutadapt v5.2	Martin[98]	N/A
HISAT2 v2.2.1	Kim et al.[99]	N/A
T2T-CHM13	National Center for Biotechnology Information (NCBI)	BioProject PRJNA559484; RefSeq assembly GCF_009914755.1
GTDB-Tk v2	Chaumeil et al.[100]	N/A
Prodigal v3.5.0	Hyatt et al.[101]	N/A
HMMER v3.3.2	Eddy[102]	N/A
mafft v7.490	Katoh et al.[103]; Kuraku et al.[104]	N/A
RAxML-ng v1.1.0	Kozlov et al.[105]	
SummarizedExperiment (R Bioconductor package)	Morgan et al.[106]	N/A
MultiAssayExperiment (R Bioconductor package)	Ramos et al.[107]	N/A
tidyverse (R package suite)	Wickham et al.[108]	N/A
janitor (R package)	Firke[109]	N/A
gt (R package)	Iannone et al.[110]	N/A
alto (R package)	Fukuyama et al.[111] https://github.com/lasy/alto	N/A
topicmodels (R package)	Grün et al.[112]	N/A
gtsummary (R package)	Sjoberg et al.[113]	N/A
ade4 (R package)	Dray and Dufour[114]	N/A
factoextra (R package)	Kassambara and Mundt[115]	N/A
DistatisR (R package)	Beaton et al.[116]	N/A
PropCIs (R package)	Scherer et al.[117]	N/A
pls (R package)	Liland et al.[118]	N/A
packMBPLSDA (R package)	Brandolini-Bunlon et al.[119]	N/A
Sickle v1.33	Joshi and Fass[120]	N/A
FastQC v0.12.1	Andrews[121]	N/A
Unicycler v0.5.1	Wick et al.[122]	N/A
Bakta database v5.1	Schwengers[123]	N/A
BUSCO v5.5	Manni[124]	N/A

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Study design and sample collection

Samples analyzed in this study were obtained as part of a previously reported phase 2b, randomized, double-blind, placebo-controlled trial of the *L. crispatus* LBP LACTIN-V (Osel, Inc., Mountain View, CA) for prevention of rBV.[37] Use of samples was approved by the Mass General Brigham Institutional Review Board (IRB Protocol #2020P002237) as well as the UCSF Institutional Review Board (IRB Protocol #19-28337). LACTIN-V is a single-strain LBP formulated as a powder containing a preservation matrix and $2x10^9$ colony-forming units per dose of the *L. crispatus* strain CTV-05, which was isolated in 1993 from a vaginal sample of a healthy US woman without BV or sexually transmitted infections in Seattle, Washington.[37,41–43] CTV-05 is administered using a pre-filled, single use vaginal applicator and was compared in the trial to a placebo formulation consisting of the preservation matrix without CTV-05. The trial was conducted at four centers within the USA and enrolled premenopausal, non-pregnant women aged 18-45 years. Eligibility criteria were previously described.[37] The original clinical trial enrolled 152 LBP and 76 placebo recipients, of whom 142 LBP and 71 placebo recipients (213 total) had samples available for our analysis. Participants' self-declared race and educational level are reported in Figure S5E, and 20% of participants identified as being of Hispanic or Latino ethnicity.

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

Briefly, potential participants attended a screening (pre-MTZ) visit at which they were determined to be eligible for the study if testing revealed presence of BV as determined by both a Nugent score of 4-10 on Gram stain of a vaginal smear[45] and presence of at least three of four Amsel criteria (characteristic vaginal discharge, >20% clue cells on microscopy of a vaginal wet prep, vaginal fluid pH >4.5, and presence of a fishy odor upon addition of 10% potassium hydroxide to a vaginal specimen),[44] as well as negative testing for HIV, syphilis, gonorrhea, chlamydia, trichomonas, and urinary tract infection. Women found to be eligible based on this evaluation completed 5 days of intravaginal MTZ therapy within 30 days of their pre-MTZ visit (Figures 1A and S1A). They then returned to the trial clinic within 48 hours of completing antibiotics (post-MTZ visit) and were randomized in a 2:1 ratio to receive either LBP or placebo after providing written informed consent. The first dose of LBP or placebo was clinician-administered at the randomization (post-MTZ) visit, then doses were vaginally self-delivered daily for the next four days, then twice weekly for ten additional weeks. In-person study visits were scheduled 4, 8, 12, and 24 weeks after randomization, at which vaginal swabs were collected and stored (details below) and clinical report forms (CRFs) were completed. In addition, two phone visits were planned at week 16 and 20 during which a subset of the clinical report forms were filled to capture information on adverse events, menstruation, concomitant medication use, and sexual behavior protected or unprotected by condoms. Participants who desired additional in-person visits were invited to present to the clinics. Swabs and CRFs were collected at these additional visits.

Clinician-collected vaginal swab samples were obtained via speculum exam at in-person study visits. Two types of swabs were collected in parallel for analysis.[37] One set of swabs was collected at all in-person visits including the pre-MTZ (screening) visit using the Starplex™ Scientific Multitrans™ Collection and Transportation System (Starplex™ Scientific S1600), which comprises a plastic-shaft Dacron™-tipped swab stored in a glass bead-containing transport medium consisting gelatin (5.0 g/L), sucrose (68.46 g/L), glutamic acid (0.70 g/L), HEPES sodium salt (3.4 g/L), modified Hank's balanced salts (9.8 g/L), sodium bicarbonate (0.35 g/L), bovine serum albumin (10 g/L) and the antimicrobial agents vancomycin (0.1 g/L), amphotericin B (2.5 mg/L), and colistin (0.015 g/L) at a pH of 7.2-7.8. The other set of swabs was collected at the post-MTZ (randomization) visit and all subsequent visits in the ESwab® Liquid Based Collection and Transport System (Copan ESwab 480C®). Samples were stored at room temperature for 1-4 hours after collection, then frozen at -80°C.

Microbiota sequencing, bacterial isolation, and cytokine analysis was performed on samples stored using the Starplex™ system. Samples were thawed on ice, vortexed at maximum speed for 5 seconds, then the swabs were removed from the transport media and media was divided into aliquots and re-frozen at -80°C. Subsequent processing was performed as described below. Measurement of bacterial load via quantitative PCR (qPCR) was performed using swabs collected via the Copan ESwab® system.

Bacterial isolations and cultivation

Bacterial isolation and cultivation were performed under anaerobic conditions at 37°C in an AS-580 anaerobic chamber (Anaerobe Systems) with an atmosphere of 5% carbon dioxide, 5% hydrogen, and 90% nitrogen (Airgas, Inc.). All culture media was pre-reduced prior to use by being placed in the anaerobic chamber overnight. Bacteria were isolated and cultivated on solid media including *Lactobacillus* MRS agar (Hardy Diagnostics, #G117), Columbia Blood Agar ("CBA", Hardy Diagnostics, #A16), or CDC Anaerobe Laked Sheep Blood Agar with Kanamycin and Vancomycin ("LKV", BD BBL™ Prepared Plated Media, #221846). Known *L. crispatus* strains and not-yet-identified isolates obtained on *Lactobacillus* MRS agar were expanded by culture in liquid media consisting of *Lactobacillus* MRS broth (BD #288130) prepared according to manufacturer instructions. Isolates obtained on CBA agar or LKV agar were expanded in liquid media consisting of either Wilkins-Chalgren Anaerobe Broth (Thermo Scientific™ Oxoid™, #CM0643B; prepared according to manufacturer instructions) or of NYCIII broth (American Type Culture Collection (ATCC) medium 1685), whichever produced better growth. NYCIII broth was prepared using a slightly modified version of the standard ATCC protocol.[35] Pre-media consisted of 4 g/L HEPES (Fisher Scientific, #BP310-500), 15 g/L proteose peptone no. 3 (BD Biosciences, #BD 211693), and 5 g/L sodium chloride in 875 ml distilled water, which was pH-adjusted to 7.3 and autoclaved on liquid protocol at 121°C for 15 minutes, then cooled and stored at 4°C. One day before use, complete NYCIII broth was prepared from the autoclaved, cooled pre-media by adding dextrose (from a stock of 3 g per 45 ml; Fisher Chemical™, #D16-500) at 7.5% v/v, yeast extract solution (Gibco, #18180-059) at 2.5% v/v, and heat-inactivated horse serum (Gibco, #26050070) at 10% v/v, then sterilized by passage through a 0.22 μm vacuum filter.

To establish a complete genome sequence for the CTV-05 strain, a cryopreserved pure culture of CTV-05 was obtained from Osel, Inc. The strain was streaked for isolation on *Lactobacillus* MRS agar and cultured for 48 hours, then a single colony was picked into *Lactobacillus* MRS broth and incubated for 20 hours. The broth culture was harvested by centrifugation to obtain bacterial pellets for genomic DNA extraction and sequencing.

Bacterial isolations were performed from twelve selected trial samples (see Figures 3C and S2A) using a modification of previously described methods.[35] Since samples had been collected into Starplex™ transport medium containing antibiotics (see above), the samples were thawed on ice, immediately diluted into pre-reduced Dulbecco's Phosphate Buffered Saline ("PBS", Millipore Sigma, #D8537) at 1:12 v/v, centrifuged for 10 minutes at 10,000 rcf, then supernatant was removed. Pellets were re-suspended in PBS and re-centrifuged with removal of supernatant two more times, then resuspended in PBS, diluted in serial 10-fold dilutions, and 100 μl aliquots from each dilution were plated evenly on MRS, LKV, and CBA agar in parallel. Plates were incubated for 7 days and multiple examples of each distinct colony morphology from each sample were picked and subcultured onto solid media of the same type as the source media, with an emphasis on colonies from MRS agar with characteristic *L. crispatus* morphology (e.g., Figure 3D). After sub-culture for 3-7 days, colonies of the sub-cultured bacteria were picked into MRS broth (for colonies isolated on MRS agar) or into both NYCIII broth and Wilkens-Chalgren broth in parallel and incubated for 1-4 days, depending on growth rate. The resulting liquid

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

cultures were then cryopreserved, with aliquots of each culture centrifuged and pellets saved for genomic DNA extraction and sequencing (see below).

METHOD DETAILS

DNA extraction for short-read sequencing

Total nucleic acids (TNA) extraction from cervicovaginal swabs for microbiota profiling was performed via a phenol-chloroform method, which includes a previously described bead beating process to disrupt bacteria[70] and modified for processing in 96-well plate format (Phenol:Chloroform:IAA, 25:24:1, pH 6.6, Invitrogen, #AM9730, which has since been discontinued; Sodium Dodecyl Sulfate 20% Solution, Fisher Scientific, #BP1311-200; EDTA, Invitrogen, #AM9260G; 2-Propanol, Sigma, # I9516-500ML; 3M Sodium Acetate, pH 5.5, Life Technologies, #AM9740). Aliquoted samples were thawed on ice, then TNA extraction was performed using 200uL of well-mixed Star media. The extracted TNA sample was eluted into 80uL of TE buffer (Promega, #V6321).

To extract bacterial genomic DNA (gDNA) for genome sequencing of cultured bacterial strains, each isolated strain was streaked on the indicated solid media and a single, clonal colony was picked into broth culture and incubated in static culture under anaerobic conditions for between 18 and 120 hours (depending on strain growth kinetics). Cultures were centrifuged and gDNA was extracted from the pellets using a plate-based protocol including a bead beating process and combining phenol-chloroform isolation[125] with QIAamp 96 DNA QIAcube HT kit (Qiagen, #51331) procedures.

Bacterial 16S rRNA gene sequencing

Bacterial microbiota taxonomic composition in cervicovaginal samples was determined by sequencing the V4 region of the bacterial 16S ribosomal RNA (rRNA) gene. The V4 region was amplified via polymerase chain reaction (PCR) using the primer set 515F/806R at 200 pM each (515F primer sequence 5'- AATGATACGGCGACCACCGAGACGTACGTACGGTGTGCCAGCMGCCGCGGTAA-3' and barcoded 806R primer sequence 5'-CAAGCAGAAGACGGCATACGAGATXXXXXXXXXXXXAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT-3', in which the underlined sequences in each primer represent the regions of complementarity to 5' and 3' ends of the V4 region of the bacterial 16S rRNA gene, respectively, and the barcode positions in the 806R primer are indicated by X; IDT), with the 806R primers barcoded for multiplexing.[35,83] PCR was performed in 25 μl reactions containing 1X Q5 reaction buffer (NEB, #B9027), 0.2 mM of dNTPs (NEB, #N0447), 0.2 μM of each primer, 0.5 unit of Q5 high-fidelity DNA polymerase (NEB, #M0491), and 2 μl of the TNA sample. PCR was performed in triplicate for each sample in the following program: 98°C for 30 seconds, followed by 30 cycles of 98°C for 10 seconds, 60°C for 30 seconds, and 72°C for 20 seconds, followed by a final extension at 72°C for 2 minutes. The triplicate PCR reactions for each sample were combined and amplicon production and size were confirmed on an agarose gel. Negative controls with PCR-quality water (Invitrogen, #10977) as a template were amplified in parallel for each primer barcode mix and assessed in parallel by gel electrophoresis to confirm absence of contamination and non-specific amplification. PCR products were pooled, with the amount for each sample semi-quantitatively adjusted based on its gel band intensity, then purified with a QIAquick PCR purification kit (Qiagen, #28104) and quality controlled with the Qubit™ 4 Fluorometer (Invitrogen #Q33226), and TapeStation (Agilent Technologies, 4200 TapeStation). Libraries were mixed with 10% PhiX and single-end sequenced on an Illumina MiSeq using a 300-cycle v2 kit (Illumina, #MS-102-2002) employing the custom Earth Microbiome Project sequencing primers (Read 1 sequencing primer sequence: 5'-ACGTACGTACGGTGTGCCAGCMGCCGCGGTAA-3'; read 2 sequencing primer sequence: 5'-ACGTACGTACCCGGACTACHVGGGTWTCTAAT-3'; index sequencing primer sequence: 5'-ATTAGAWACCCBDGTAGTCCGGCTGACTGACT-3'; IDT).[83] Negative controls for TNA extractions and PCRs were included in each sequencing library. Study samples were sequenced in a total of six libraries. Samples with read counts < 10,000 in initial libraries were re-amplified and re-pooled into subsequent libraries and data from the run producing the highest number of reads for each sample (if the sample was sequenced multiple times) were selected for subsequent analysis. Analyzable 16S rRNA gene data was generated from a total of 1152 out of 1156 available trial samples, with the remaining 4 samples failing due to technical challenges with extraction, amplification, or sequencing ($<10^3$ processed reads per sample).

Bacterial short-read genome sequencing

Shotgun metagenomic and genomic libraries were prepared following a modified protocol of Baym et al.,[126] using the Nextera DNA Library Preparation Kit (Illumina, #20034211) and KAPA HiFi Library Amplification Kit (Kapa Biosystems, #KK2602). In brief, DNA from each sample was standardized to a concentration of 1ng/mL after quantification with SYBR Green (Invitrogen, #S7653), followed by simultaneous fragmentation and sequencing adaptor incorporation by mixing 1ng of DNA (1 mL) with 1.25 ml TD buffer and 0.25 mL TDE1 provided in the Nextera kit and incubating for 9min at 55°C. Tagmented DNA fragments were amplified in PCR using the KAPA high fidelity library amplification reagents, with Illumina adaptor sequences and sample barcodes incorporated in primers. PCR products were pooled, purified with magnetic beads (MagBio Genomics #AC-60050) and paired-end sequenced on Illumina NovaSeq X with a 300-cycle kit (Psomagen, Inc.).

Bacterial 16S rRNA gene sequence annotation

Demultiplexing of Illumina MiSeq bacterial 16S rRNA gene sequence data was performed using QIIME 1 version 1.9.188.[84] Mapping files created in QIIME 1 format were validated using validate_mapping_file.py, then sequences were demultiplexed with split_libraries_fastq.py using parameter store_demultiplexed_fastq and no quality filtering or trimming, and demultiplexed sequences were

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

organized into individual fastq files using split_sequence_file_on_sample_ids.py. Sequence reads were trimmed and filtered using dada2 version 1.6.0,[85] trimming at positions 10 (left) and 230 (right) using the filterAndTrim function with truncQ = 11, MaxEE = 2, and MaxN = 0. Sequences were then inferred, then initial taxonomy assigned using the dada2 assignTaxonomy function with the RDP training database rdp_train_set_16.fa.gz (https://www.mothur.org/wiki/RDP_reference_files). Amplicon sequence variant (ASV) taxonomic assignments were refined via extensive manual review. Final ASV taxonomic assignments are included in Data S1 for ease of review and comparison to other literature. The resulting annotated sequences were analyzed in R using phyloseq version 1.30.0[86] and custom R scripts (see STAR Methods). Sequence processing and taxonomy assignment was performed blinded to information about participants' and samples' clinical and demographic characteristics, treatments, and trial outcomes.

Total bacterial load via qPCR

Quantification of total bacterial load via qPCR was performed and reported as part of the initial LACTIN-V clinical trial.[36] In brief, DNA extracted from samples stored in the Copan ESwab® system were amplified using bacterial 16S rRNA gene primers targeting total bacteria (16S ribosomal DNA, AGAGTTTGATCCTGGCTCAG, GCTGCCTCCCGTAGGAGT, 312bp). Bacterial concentration was calculated using a standard curve based on serial dilutions of the CTV-05 strain as previously described.[37]

Assessing baseline balance between arms

Since pre-intervention (*i.e.,* pre-MTZ or post-MTZ) microbiota composition may be associated with differential microbiota composition post-intervention, imbalances between arms could lead to biases in our primary outcome benefit ratio estimates. To assess whether significant imbalances in microbiota composition existed in this cohort, we performed a PERMANOVA analysis, as implemented in the vegan R package,[127] to test whether the intervention arm (LBP *vs* placebo) explained significant variability in microbiota β-diversity, computed using the Bray-Curtis (BC) dissimilarity on ASV relative abundances.

Calculating benefit ratios

Benefit ratios, and associated confidence intervals and *p*-value were computed using Wald's method as implemented in the epitools R package.[87]

Identification of microbiota topics

Microbiota topics were identified using a modification of a previously described approach.[50] One main difference is that here, *Lactobacillus* topics (*i.e.,* topics composed exclusively of *Lactobacillus* species) were defined independently from non-*Lactobacillus* topics (*i.e.,* topics composed exclusively of non-*Lactobacillus* species), which were identified by fitting a Latent Dirichlet Allocation model[53] to the non-*Lactobacillus* ASV counts aggregated at the species level. This was done to facilitate the interpretation of topic composition and, specifically, to allow for "pure" topics for the most prevalent *Lactobacillus* species in this cohort (*L. crispatus*, *L. iners*, and *L. jensenii*).

To determine K, the optimal number of non-*Lactobacillus* topics, we relied on a method called "topic alignment"[128] which examines the robustness of topics across resolutions (increasing values of K) and provides diagnostics scores that facilitate the identification of spurious topics. We selected K to minimize the number of spurious topics (flagged by low coherence scores) and such that the number of paths (collection of similar topics across resolution) in the alignment presented a plateau[128] (Figures S1D and S1E). The estimated proportions (relative abundances) of non-*Lactobacillus* topics in each sample ($\widehat{\rho}_{ki}$ where k is the topic and i is the sample) were computed by multiplying the proportions estimated by the model on non-*Lactobacillus* counts ($\widehat{\pi}_{ki}$) by the total non-*Lactobacillus* proportions in each samples ($\widehat{\Pi}_i = \sum_{v \in V} p_{vi}$ where p_{vi} is the observed proportion of ASV v in sample i and V is the set of non-*Lactobacillus* ASVs, such that $\widehat{\rho}_{ki} = \widehat{\pi}_{ki}\ \widehat{\Pi}_i$).

Lactobacillus topics were defined as follows: *Lactobacillus* species which reached 50% of a microbiota composition in at least 10 samples made up their own topic, while the remaining *Lactobacillus* species were grouped into a single topic ("Other *L.*") as their total prevalence and abundance was overall small (Figure S1F). The composition of this topic was estimated from the species average prevalences in this cohort. The estimation of proportions of *Lactobacillus* topics in each sample was straightforward: they were computed from the proportions of the corresponding species in each sample.

Isolate genome assemblies from short-reads

Paired-end short-read genomic sequencing data from bacterial isolates were processed through a quality control and assembly pipeline. Raw reads were first trimmed for adapter sequences using Cutadapt v5.2[98] with the Nextera adapter sequence (CTGTCTCTTAT). Quality filtering was then performed using Sickle-trim v1.33[120] with a quality threshold of Q20 and minimum read length of 50 bp after trimming. Read quality was assessed using FastQC.[121] De novo genome assembly was performed using Unicycler v0.5.0 in standard mode with default parameters.[122] Unicycler produces high-quality assemblies from short reads alone by optimizing SPAdes assembly, followed by graph simplification. The resulting assemblies were annotated using Bakta v1.9.4[90] with the Bakta database v5,[123] which provides comprehensive bacterial genome annotation including coding sequences, rRNAs, tRNAs, and other genomic features. Assembly completeness and quality were evaluated using BUSCO v5.5[124] with the Bacteroidales lineage dataset (bacteroidales_odb10), providing a measure of genome completeness based on conserved single-copy orthologs. Isolate genome-assembly statistics and NCBI BioSample numbers are in Data S3; the median N50 was >23,000, median completeness >99.8%, and median contamination <0.36%.

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

CTV-05 sequencing and genome assembly

For long-read (Oxford Nanopore Technologies) genome sequencing of the CTV-05 strain, a bacterial pellet was prepared from a clonal colony cultured in *Lactobacillus* MRS broth media as described above. gDNA was extracted using a non-bead-beating protocol to minimize shearing of DNA fragments, with extractions performed and sequenced using five parallel aliquots of the same original broth culture to maximize yield and consistency. Since *L. crispatus* is known to have a thick cell wall containing surface-layer (S-layer) proteins, cell pellets were resuspended with 5M lithium chloride (Molecular Dimensions #MD2-100-43) and then washed with PBS to begin to degrade the cell wall. Using the manufacturer's protocol for the MasterPureTM Gram Positive DNA Purification Kit (Biosearch Technologies #MGP04100), the bacterial cells were then lysed, proteins and RNA were digested, and genomic DNA (gDNA) was extracted. Extracted gDNA concentration was diluted 1:5 and measured using the Qubit™ 4 Fluorometer, then DNA libraries were prepared using Oxford Nanopore Sequencing's Rapid Barcoding Kit (Oxford Nanopore Technologies #SQK-RBK004; note that this product has been discontinued as of March 2024). In brief, 400 ng of gDNA from five replicate extractions was barcoded, pooled, purified, and then loaded onto the MinION Flow Cell using R9.4.1 chemistry (Oxford Nanopore Technologies #FLO-MIN106D; note that this product has been discontinued as of July 2024) in the MinION Sequencing Device (Oxford Nanopore Technologies #MIN-101B). To produce the maximum number of reads possible using the Rapid Barcoding Kit, the sequencing run was continued for a full 72 hours.

Demultiplexed genomic long-reads from MinKNOW v2.2 were concatenated, and reads were separately filtered for quality and length using Filtlong v0.2.1 (https://github.com/rrwick/Filtlong). Filtered reads were subsampled to make 12 different read sets using Trycycler v0.5.3[91] and duplicate reads from each read set were removed using BBMap v39.00.[92] The read sets were assembled using Flye v2.9.1,[93] Minipolish v0.1.3,[94] Raven v1.8.1,[95] and Canu v2.2.[96] Contigs from different assembly methods were clustered based on similarity to generate consensus sequences and improve assembly quality using Trycycler. Clusters were manually chosen and Trycycler was used to reconcile the contigs within each cluster. Multiple sequence alignment and read partitioning were performed on the reconciled contigs using Trycycler to understand sequence variation within each cluster and assign each read to the cluster it best aligned with respectively. Trycycler was then used to generate a consensus contig sequence for each cluster based on the multiple sequence alignment and read partitioning steps. Medaka v1.7.2 (https://github.com/nanoporetech/medaka), a tool for creating consensus sequences specifically from nanopore sequence data, was used to polish the Trycycler consensus sequences. The consensus sequences were then combined, after which Polypolish v0.5.0[97] was used to polish the long-read sequence with Illumina short-read data generated from an aliquot of the same clonal culture of CTV-05 (extracted and sequenced as described above). The complete CTV-05 genome was annotated using Bakta v1.7.0[90] and published online as part of this manuscript under NCBI BioSample: SAMN55039759, which includes the unassembled short- and long-reads described above as well as the complete CTV-05 genome assembly. BioSample reference number SAMN55039759 is included under the study's overall BioProject: PRJNA1303956.

L. crispatus SNV database for strain analysis

A GT-Pro database of biallelic core sites was constructed using 236 *L. crispatus* genomes using k-mers with thresholds for SNV prevalence of 0.95 and minor allele frequencies of 0.02 with MAAST.[55,88,89] Then k-mers for samples were counted to generate a metagenotype, or a matrix of the number of reads at biallelic SNV sites for a single species.[54] This matrix was then used by StrainFacts to infer suitable allele and relative abundance matrices that represent strains across all given samples.[55]

L. crispatus strain inference in metagenomes

Raw metagenomic paired-end reads from vaginal swab samples were quality-trimmed and adapter-filtered using Cutadapt v3.9.[98] Human reads were depleted by aligning reads to the T2T-CHM13 reference genome using HISAT2 v2.2.1.[99] Samples were sequenced to a median depth of 22,329,925 (IQR: 14,928,888–33,699,201) reads, yielding a median of 2,748,850 (IQR: 1,186,960–7,950,264) analyzable microbial reads after host and quality filtering for a median of 12.7% (IQR: 4.53–25.9%) of reads retained after processing (sample-level read depths pre- and post-processing, as well as NCBI BioSample numbers, are provided in Data S2). Processed reads were profiled using GT-Pro (v1.0.1) against our custom SNV database consisting of ~26,000 biallelic SNV sites.[88] Resulting metagenotype profiles were analyzed with StrainFacts (v0.6.0) to infer strain genotypes and abundances for *L. crispatus*, excluding samples with <25% coverage of the biallelic SNVs.[54] The fit was then run through a cleanup step to remove low abundance strains and strains with extremely similar genotypes using sfacts cleanup_fit –abundance 0.01 –dissimilarity 0.01 –discretized. Analysis was performed on samples with ≥5% overall relative abundance of *L. crispatus* as determined by 16S rRNA gene sequencing based on results of a separate validation study evaluating parameters for accurate StrainFacts LBP strain inference.[55] Since most samples had *L. crispatus* relative abundance either substantially higher or lower than 5%, with only 31 (<3%) exhibiting relative abundance between 2.5-5% (e.g., Figure 1D), decreasing this threshold would likely have increased error while negligibly increasing the number of samples selected for analysis. *L. crispatus* strain inference was successful in 313 of 343 samples in which species relative abundance exceeded the analysis threshold, while strain analysis was unsuccessful in the remaining 30 samples due to failed or inadequately deep shotgun metagenomic sequencing. A 10% strain proportional abundance was used as a threshold for confident strain identification based on prior simulations and experimental validation.[55] Most samples had *L. crispatus* fractional strain abundances either substantially higher or lower than 10%, with only 12 of 313 samples (3.8%) containing inferred strains with fractional abundance between 7.5% and 12.5%, showing strain detection was robust to threshold choice.

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

Strain analysis of isolate genomes

Isolate genomes were taxonomically identified using GTDB-Tk v2.[100] Isolate taxonomic details are in Data S3. Gene-calling was performed for *L. crispatus* isolate genomes, including the completed CTV-05 genome (see above) using Prodigal.[101] Core ribosomal proteins were identified using HMMER,[102] aligned using mafft,[103,104] then concatenated and a phylogenetic tree was constructed using RAxML-NG,[105] which was used to compute branch lengths to identify isolates representing distinct strains and determine which isolates represented CTV-05 or endogenous strains. The tree was rooted to *L. jensenii*. GT-Pro[88] was then used to calculate genotypes from representative isolate reads for comparison to metagenomically inferred strain genotypes, calculated based on Jaccard similarity of the StrainFacts genotypes.[55]

Mucosal cytokine and chemokine measurement

Cytokines and chemokines were measured using a previously described custom 20-plex High Sensitivity Luminex Assay Kit from EMD Millipore that measures interferon gamma-induced protein (IP-10), interleukin 8 (IL-8), interleukin 6 (IL-6), monokine induced by interferon gamma (MIG), interferon-inducible T cell alpha chemoattractant (ITAC), interleukin 1 alpha (IL-1α), interleukin 1 beta (IL-1β), macrophage inflammatory protein-1 alpha (MIP-1α), macrophage inflammatory protein-1 beta (MIP-1β), macrophage inflammatory protein-3 alpha (MIP-3α), tumour necrosis factor alpha (TNFα), interleukin 21 (IL-21), interleukin 17 (IL-17), interferon gamma (IFNγ), interleukin 23 (IL-23), interleukin 12 (IL-12 p70), interleukin 13 (IL-13), interleukin 10 (IL-10), interleukin 4 (IL-4), and interleukin 5 (IL-5).[6] Reagents were used as supplied by the manufacturer, except the Mixed Beads solution, detection antibody mixture, and Streptavidin-phycoerythrin (Strep-PE) were diluted 3-fold (1:2 volume:volume ratio) for use. The Mixed Beads were diluted in Bead Diluent, while the detection antibodies and Strep-PE were diluted individually in Assay Buffer. All reagents were from a single manufacturer lot to minimize batch effects.

Samples were first pre-processed to remove interfering mucus and debris. In brief, 200 μL aliquots of each vaginal swab supernatant were thawed on ice, vortexed for 10 seconds to resuspend, then centrifuged at 1000 rcf for 15 minutes at 4°C to pellet mucous and cells. The maximum amount of supernatant recoverable from each sample without disturbing the resulting pellet was then transferred to a well of a 0.22 m PVDF filter plate (EMD Millipore #MSGVS2210), filtered by centrifugation at 2,451 rcf for 1 hour at 4°C, then transferred to freezer-safe tubes and stored at -80°C until assayed. Samples were assayed in 96-well plates as per manufacturer protocol with the addition of periodic sonication steps to minimize bead clumping. Each plate assay included one blank background control, 7 standard serial dilutions assayed in duplicate, two manufacturer-supplied Quality Control (QC) samples assayed in duplicate, 72 experimental samples, and 4 biological control samples used as internal quality controls on all plates. Batches of standards and QC samples were prepared from the kit on the day of the assay according to manufacturer protocol. For each sample or standard, 25 μL each of Assay Buffer, Mixed Beads, and sample were aliquoted into each plate well. After the addition of Mixed Beads, all incubation steps were performed while the plate was protected from light. Plates were sonicated for 30 seconds at room temperature in a bath sonicator, then incubated for 16-18 hours on a horizontal plate shaker at 600 rpm in 4°C. An additional 30 second sonication at room temperature was then performed. A handheld plate magnet was applied to the base of the plate to retain the magnetic beads and excess sample and reagents were removed via three consecutive washes using manufacturer-supplied 1X Wash Buffer. Detection antibodies were added as per manufacturer protocol and plates were incubated for 1 hour at room temperature on a horizontal plate shaker at 600 rpm. Next, Strep-PE reagent was added, followed by a 30 minute incubation at room temperature on a horizontal plate shaker at 600 rpm. Plates were washed 3 times as above using the magnet. After the final wash, 150 μL of sheath fluid was added to all wells and a final 30 second sonication at room temperature was performed. Analyte concentrations were measured using a Luminex® FLEXMAP 3D instrument with Luminex xPONENT® software (Diasorin).

Cytokine concentration transformations

Cytokine and chemokine concentrations were log-transformed (log) to stabilize the mean-variance relationship. Compounds with concentration values below the lower limit of quantification (LLOQ) were imputed at half the LLOQ; those with concentration values above the upper limit of quantification (ULOQ) were imputed at the ULOQ (Figure S3A).

Measurement of soluble molecules in swab samples can be complicated by a "size effect" phenomenon whereby a component of between-swab variation in concentrations is due to technical differences in the amount of material collected on each swab. In such situations, the first principal component (PC1), which reflects the "size" of observations,[56] can primarily be driven by the amount of material on the swab rather than biological variation. Consistent with this size effect phenomenon, we observed that per-sample cytokine/chemokine concentrations were highly collinear, with PC1 of the standardized log-transformed cytokine concentrations accounting for >50% of the total variance (Figure S3B) but showed no substantial correlation with clinical and biological factors such as proportions of total *Lactobacillus* or of *Lactobacillus crispatus*, or participants' contraceptives. Such size effects can be addressed by subtracting PC1 from the data. We therefore performed PC1 subtraction by re-assigning scores corresponding to the 1st PC to a value of 0, then transforming the data back to its original variable space using the transposed rotation matrix. Sensitivity analyses employing unadjusted concentrations revealed qualitatively similar findings to results obtained using the PC1-subtracted data, but with reduced magnitude. Finally, due to the large uncertainty regarding their distributions, cytokines with values below the LLOQ in ≥60% of the samples or with values above the ULOQ in ≥30% of the samples were excluded from the analysis (Figure S3A).

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

Association of microbiota and cytokines

To quantify the overall association between microbiota composition and cytokine profiles, we computed the RV coefficient between the microbiota composition as expressed as topic proportions and cytokine transformed $\log_{10}$-concentrations. Associated *p*-value was computed with a permutation test.[114,129] To confirm that correlation between these tables was not driven by potential subject (longitudinal) effects, we also computed correlation between the two tables at each visit independently. If correlations were found significant, further analyses were carried out to characterize the relationship between tables. Specifically, DISTATIS,[59,130] as implemented in the distatisR R package, was used. This method presents the advantage of estimating compromise scores (and partial residual scores) directly from dissimilarity matrices such that ecological measures of (dis)similarity, such as the Bray-Curtis dissimilarity, can be used for representing microbiota composition. Associations with the original variables can then be assessed using correlations and displayed in correlation circles as typically done with PCA, PCoA, or NMDS results. The minimum number of components used for display was chosen based on the presence of an elbow in the DiSTATIS scree plot.

Estimating effect heterogeneity

Binary outcomes (Y: ≥50% colonization by *L. crispatus* at week 12 or 24 or absence of rBV by week 12 or 24) were predicted using logistic regression (logit link function) with input variables representing the intervention arm (*A*: LBP or placebo), the baseline (pre-MTZ) microbiota (*V*: stratified by most prevalent genus, CST, or topic proportions; Figures 5C–5F and S4) or self-declared racial group (Figures 6E and 6F), and their interactions (*A:V*). In stratification analyses, strata with less than two participants per arm were excluded from the analysis and visualizations. Since *Lactobacillus* species other than *L. iners* had low pre-MTZ relative abundance, prior to determining fits relying on topic proportions, *Lactobacillus*-dominated topics were agglomerated such that the proportion of total *Lactobacillus* was considered, which improved fit convergence. The statistical significance of quantitative effect heterogeneity was tested by comparing the null model (which only includes the intervention arm as predictor) with the full model where pre-MTZ microbiota (or self-declared racial group) and interaction effects were included (analysis of deviance F test). *P*-values were adjusted to account for multiple hypotheses testing using the Benjamini-Hochberg procedure to control the false discovery rate.[131] For the stratified analyses, CI were computed using Wilson scores due to the small sample size in each stratum. When pre-MTZ microbiota composition or self-declared racial group was included as a quantitative multivariate variable, counterfactual probabilities of success and associated 95% confidence intervals were predicted from the logistic regression fitted model by setting the intervention arm to LBP or placebo. The predicted participant-level odd ratios were computed as the ratio between predicted odds of *L. crispatus* colonization at week 12 and 24 or rBV by week 12 or 24 had participants been receiving LBP or the placebo.

Factors predictive of microbiota category

To identify demographic, clinical, behavioral, or microbiologic factors associated with successful *L. crispatus* colonization (≥ 50%) in the intervention arm, we relied on a multiblock partial least square discriminant analysis (MB-PLS-DA)[62] where the 3-category response variable indicated whether (1) relative abundance of *L. crispatus* was ≥50%, (2) relative abundance of total *Lactobacillus* (any species) was ≥50% but relative abundance of *L. crispatus* was <50%, or (3) the relative abundance of total *Lactobacillus* was <50%. MB-PLS-DA relies on the same principles as PLS-DA but allows for explanatory variables to be grouped into several thematic blocks, which enables the computation of block importance indices and covariances with the response block.[61,62] Our analysis grouped explanatory variables into 13 blocks (Figure 6A; Data S4). The first block characterized participant demographics, blocks 2-4 described participant baseline vaginal ecosystem (*i.e.*, pre-MTZ microbiota composition and diversity, pH, cytokines), blocks 5-8 quantified the vaginal ecosystem at the previous visit (*e.g.,* at the post-MTZ visit for the initial phase), while blocks 9-13 characterized sexual behavior, douching/bleeding, antibiotic use, and product adherence (Data S4).

Several explanatory variables were correlated. For example, the abundances of some cytokines correlated with microbiota composition (Figure 4), and the baseline α-diversity was lower in White participants (Figure S5F, p-value < 0.05). We addressed these existing correlations between explanatory blocks and/or variables in two different ways depending on our assumptions on the underlying correlation source or cause. We either relied on nested models to evaluate the additive predictive power of specific blocks (*e.g.,* demographics) that had variables correlated with microbiological blocks or used one variable or one block to predict the values of another one and included the residuals instead of the observed values in the model. This was indicated by the symbol (r) in the variable or block names. We used this approach for the cytokine blocks whose residuals were computed using their PLS-predicted abundances based on microbiota composition. Similarly, we computed residual microbiota composition, α-diversity, and pH at the previous visit such that topic proportions, α-diversity, or pH were relative to those expected based on the participants' colonization status at the same visit.

Categorical variables such as self-declared race or birth control were included using one-hot encoding, and variables were standardized. To ensure convergence of the fits in cross-validation or using the bootstrap, we added Gaussian noise with very small variance ($\sigma_i^2 = 10^{-3} S_i^2$ where σ_i^2 is the variance of the Gaussian noise for variable i, and S_i^2 the empirical variance of the i^{th} variable) to one-hot encoded variables that had few participants in some categories.

Given that our categorical response had three categories, two latent components for the MB-PLS-DA models were a natural choice to avoid overfitting. Further, we performed cross-validation analyses (40 random 75-25% split into calibration and validation sets) and found that two latent components maximized the mean average F1 score on the validation sets for all models except for the initial phase of the placebo in which no variables were found to predict our response, leading to poor and unstable performances in cross-validation (Figures S5B and S6A). This was consistent with the presence of an elbow in the screeplots and robust to other

Please cite this article in press as: Bloom et al., Vaginal microbiota impacts of a Lactobacillus crispatus live biotherapeutic and predictors of colonization in randomized controlled trial, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.03.003

choices of metrics such as accuracy or root mean square error (RMSE). Overall, results on calibration and validation sets indicated initial-phase colonization was more difficult to explain (average validation F1 score: 0.46, 95%CI 0.42-0.49) and explained by more factors (best validation F1 score obtained for the full model) than colonization at later visits (best average validation F1 obtained for the minimal models of 0.64; 95%CI 0.62-0.66 for continuation phase and 0.47; 95%CI 0.46-0.49 follow-up phase, Figures 6A, S5B, and S5C). Directions of associations between explanatory variables and outcomes were consistent across nested models. Relative cumulative block importance indices were computed by dividing cumulative block importance, the BIPC as in Brandolini-Bunlon et al.,[62] by the total inertia of each block, which corresponds to the expected BIPC if all variables had a similar importance; 95% CIs were calculated using a bootstrap approach.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

Statistical analyses are described in figure legends as appropriate and details for individual analyses are provided in the corresponding sections or the STAR Methods describing the relevant bioinformatic and experimental method details.

Software

R packages used in the analyses included SummarizedExperiment[106], MultiAssayExperiment[107], tidyverse,[108] janitor,[109] gt,[110] alto,[111] topicmodels,[112] gtsummary,[113] ade4,[114] factoextra,[115] DistatisR,[116] PropCIs,[117] pls,[118] packMBPLSDA[119] (see also key resources table and publicly available code).