Cervicovaginal microbiome alters transcriptomic and epigenomic signatures across cervicovaginal epithelial barriers

Lauren Anton

lanton@pennmedicine.upenn.edu

University of Pennsylvania

Ana G. Cristancho
Children's Hospital of Philadelphia, University of Pennsylvania

Briana Ferguson
University of Pennsylvania

Jacques Ravel
University of Maryland School of Medicine

Michal A. Elovitz
Icahn School of Medicine at Mount Sinai

Research Article

Abstract

Background

The cervicovaginal microbiome plays a critical role in women's health, with microbial communities dominated by *Lactobacillus* species considered optimal. In contrast, the depletion of lactobacilli and the presence of a diverse array of strict and facultative anaerobes, such as *Gardnerella vaginalis*, have been linked with adverse reproductive outcomes. Despite these associations, the molecular mechanisms by which host-microbial interactions modulate cervical and vaginal epithelial function remains poorly understood.

Results

In this study, we used RNA sequencing to characterize the transcriptional response of cervicovaginal epithelial cells exposed to the culture supernatants of common vaginal bacteria. Our findings revealed that *G. vaginalis* culture supernatants upregulate genes associated with an activated innate immune response and increased cell death. Conversely, *Lactobacillus crispatus* culture supernatants induced transcriptional changes indicative of epigenomic modeling in ectocervical epithelial cells. Epigenomic modification by *L. crispatus*, was confirmed by ATAC-sequencing, which demonstrated reduced chromatin accessibility.

Conclusions

These results provide new insights into host-microbe interactions within the lower reproductive tract and suggests that modulating the vaginal microbiome could offer innovative therapeutic strategies to improve reproductive health.

Introduction

The female lower reproductive tract is a complex ecosystem comprised of host epithelial and immune cells, a microbiome and a complex mixture of metabolites [1]. Over the past decade, the cervicovaginal microbiome has become the focus of extensive research due to its intricate and integral role in reproductive health and disease. High throughput 16S rRNA gene amplicon sequencing has allowed detailed characterization of vaginal microbiota composition in both pregnant and non-pregnant individuals [2–4]. Traditionally, the vaginal microbiome has been defined by the presence or absence of *Lactobacillus* species [2, 5, 6]. Cervicovaginal microbiomes dominated by *Lactobacillus* species are generally considered optimal and are associated with positive reproductive health outcomes. In contrast, microbiomes lacking lactobacilli and comprising a wide array of strict and facultative anaerobes have been linked to a range of adverse gynecological and reproductive outcomes including infertility [7, 8], sexually transmitted infections (STIs) (*e.g.*, human papilloma virus [HPV] [9] and human immunodeficiency virus [HIV] [10]), and pregnancy complications such as preterm birth [11, 12]. While the cervicovaginal microbiome is less taxonomically diverse than those at some body sites such as the

gut, the presence of con-specific genotypes cohabitating within these microbiomes adds to their complexity [13]. Importantly, not all women with lactobacilli-deficient, anaerobe-rich cervicovaginal microbiomes experience negative clinical outcomes. This fact suggests that the contribution of a suboptimal cervicovaginal microbiome to adverse reproductive outcomes may depend on host-microbe or microbe-microbe interactions within the cervicovaginal space. Understanding the complexity of these interactions is essential to elucidate the precise mechanisms by which vaginal bacteria modulate host epithelial functions and contribute to adverse health outcomes.

The vaginal microbiome can interact with all epithelial barriers in the cervicovaginal space. These epithelial barriers are unique because the cells lining this space have distinct embryological origins, resulting in specialized cell functions, such as mucus production in the cervix [14–17]. As the primary site of entry for pathogens, the integrity of this barrier is critical, and its disruption is associated with increased susceptibility to STIs (*e.g.*, chlamydia, gonorrhea, HIV) [18, 19]. Metabolites, proteins and other products from common vaginal bacteria, including *Gardnerella vaginalis* and *Lactobacillus crispatus*, have been shown to exert distinct biological effects in the cervicovaginal space. Culture supernatants from non-optimal bacteria, such as *G. vaginalis*, trigger innate immune responses in cervicovaginal epithelial cells, a process thought to protect the epithelial barrier [20–23]. In contrast, lactic acid, a metabolite produced by *Lactobacillus* species, helps maintain an acidic vaginal pH [24], exhibits significant anti-microbial and anti-viral activity against Bacterial Vaginosis (BV)-associated bacteria [25, 26] and HIV [27, 28], modulates inflammatory responses [29–31] and increase cervical epithelial barrier integrity [32, 33]. Despite these findings, the molecular mechanisms underlying the beneficial and/or harmful effects of these vaginal bacteria remain poorly understood.

Elucidating the molecular mechanisms by which cervicovaginal microbiomes modulate host responses in the lower reproductive tract is critical to understanding their role in reproductive health and disease. The objectives of this study were to: 1) use unbiased discovery-based RNA-sequencing to identify genes and functional pathways in cervical and vaginal epithelial cells altered by exposure to *G. vaginalis* or *L. crispatus* culture supernatants; 2) characterize the immune pathways activated by *G. vaginalis* culture supernatants; and 3) uncover the molecular mechanisms by which *L. crispatus* culture supernatants optimize cervical and vaginal epithelial barriers.

Results

Cervicovaginal epithelial cell gene transcription and associated functional pathways are differentially modulated by L. crispatus and G. vaginalis culture supernatants

Exposure of cervical and vaginal epithelial cells to *L. crispatus* or *G. vaginalis* bacteria-free culture supernatants resulted in significant differences in gene expression profiles. Principal component analysis (PCA) plots revealed distinct clustering of gene expression profiles by bacterial exposure across ectocervical, endocervical and vaginal epithelial cells (Fig. 1A-C). Notably, exposure to *L. crispatus* culture supernatant showed the most pronounced separation from the NYCIII bacterial culture medium

control (Fig. 1A-C). For *G. vaginalis* culture supernatants, clustering patterns were distinct from the NYCIII control in endocervical and vaginal cells but not in ectocervical cells, highlighting cell type-specific responses (Fig. 1A-C). To ensure robust identification of differentially expressed genes, we applied stringent criteria: an adjusted p-value ≤ 0.05 and a log2 fold change ≥ 1 or ≤ -1 (Fig. 1D, Supplemental Tables 1 and 2). Using these thresholds, we found that the number of differentially expressed genes was highest after exposure to *L. crispatus* culture supernatants, followed by *G. vaginalis* (Fig. 1D). A minority of these genes overlapped between cell types for each bacterial culture supernatant exposure (Fig. 1E). For *G. vaginalis*, endocervical and vaginal epithelial cells exhibited the highest number of cell-specific differentially expressed genes, while *L. crispatus* culture supernatant elicited the highest number of cell-specific genes in ectocervical and endocervical epithelial cells (Fig. 1E, Supplemental Table 3). Within each cell type, we performed comparisons to identify commonly modulated gene expression across culture supernatant exposures (Fig. 1F-H, Supplemental Table 4). This analysis revealed unique genes differentially regulated by *G. vaginalis* and *L. crispatus* culture supernatants, demonstrating specific molecular effects of supernatant exposures on cervicovaginal epithelial cell transcription (Supplemental Table 5).

We conducted gene ontology (GO) analysis of upregulated and downregulated genes for each cell line and culture supernatant exposure combination to investigate these transcriptional differences further. This analysis uncovered overlapping or distinct cellular responses to bacterial culture supernatant exposures (Supplemental Tables 6a-i) [34, 35]. An aggregate dysregulation score was calculated for all affected pathways per sample, and averages for each cell line were visualized as heatmaps to reflect the diversity of gene expression changes between bacterial culture supernatant exposures (Fig. 2A) [34]. Unsupervised clustering of GO terms revealed specific trends, and a word cloud was generated to highlight the top GO terms associated with each cluster [36, 37]. Notably, themes of inflammatory and transcriptional dysregulation emerged. To identify the most critically dysregulated pathways in each cell type, we clustered GO term differences by cell line and bacterial supernatant exposure, focusing on the top clusters for further investigation [34]. Genes upregulated by *G. vaginalis* culture supernatants were predominantly associated with inflammation functional pathways (Fig. 2B-D). In contrast, exposure to *L. crispatus* culture supernatant was associated with modulation of transcriptional pathways, including histone modifications, RNA polymerase II and DNA binding (Fig. 2A, 2E-G).

G. vaginalis, but not L. crispatus, culture supernatant dysregulates the innate immune response

Exposure of cervicovaginal cells to *G. vaginalis* culture supernatants led to the differential expression of genes significantly associated with innate inflammation-related functional pathways (Fig. 2A, 2B-2D). Canonical NF-kB pathway genes encoding for multiple chemokines and cytokines such as IL-8, IL-6 and TNFα (Supplemental Table 2) were upregulated, consistent with prior findings by our group and others [22, 38–40]. Additionally, several anti-microbial peptides (AMPs), key components of the innate immune response, were upregulated in cervicovaginal epithelial cells following exposure to *G. vaginalis* culture supernatant. These included Chemokine Ligand 20 (CCL20), Secretory Leukocyte Peptidase Inhibitor (SLPI), Lipocalin 2 (LCN2) and S100 Calcium Binding Protein 8 (S100A8/A9, Calgranulin). Of these,

CCL20 was the only gene consistently upregulated across all three cell lines (adjusted $p < 0.05$, Fig. 3A, 3C, 3E, Supplemental Table 7). In ectocervical and endocervical cells, all four AMP genes were significantly upregulated following exposure to *G. vaginalis* culture supernatant (adjusted $p < 0.05$, Fig. 3A and C, Supplemental Table 7). However, in vaginal epithelial cells, only CCL20 was upregulated under the same conditions (adjusted $p < 0.05$, Fig. 3E, Supplemental Table 7).

Protein-level analysis via ELISA confirmed the overexpression of CCL20 and S100A8 (both $p < 0.05$) after exposure to *G. vaginalis* culture supernatants. In contrast, protein levels of SLPI or LCN2 remained unchanged despite their transcriptional upregulation (Fig. 3B, 3D, 3F). Furthermore, *G. vaginalis* culture supernatants increased cell death in ectocervical and endocervical cells but not vaginal epithelial cells. In comparison, *L. crispatus* culture supernatant had no effect on cell death (Supplemental Fig. 1).

Exposure to *L. crispatus* culture supernatants showed distinct effects on AMP gene expression. The gene expression of S100A8 was significantly reduced across all three cell lines ($p < 0.001$, Fig. 3A, 3C, 3E, Supplemental Table 7). Additionally, LCN2 was downregulated in endocervical and vaginal epithelial cells ($p < 0.0001$, Fig. 3C and 3E, Supplemental Table 7), while CCL20 gene expression was decreased specifically in endocervical cells ($p < 0.0001$, Fig. 3C, Supplemental Table 7). In ectocervical cells, however, SLPI and LCN2 gene expression were increased ($p < 0.0001$, Fig. 3A, Supplemental Table 7). Despite these transcriptional changes, exposure to *L. crispatus* culture supernatant was not associated with consistent changes in AMP protein levels (Fig. 3B, 3D and 3F) across cell lines.

L. crispatus culture supernatants alter chromatin accessibility

Exposure to *L. crispatus* culture supernatants resulted in significant changes in the expression of genes related to histone and transcriptional regulation in cervical and vaginal epithelial cells (Fig. 2A, 2E-G). To investigate whether these gene expression changes were associated with alterations in chromatin accessibility, we performed assay for transposase-accessible chromatin high throughput sequencing (ATAC-seq) [41–43].

The number and percentage of aligned and unaligned sequence reads were consistent across all three cell types (Supplemental Fig. 2). However, ectocervical epithelial cells exhibited lower transcription start site (TSS) enrichment scores despite showing similar quality control characteristics in alignment quality (Supplemental Fig. 3). A strong correlation in normalized sequence read counts between conditions was observed for each cell type, indicating high-quality samples (Supplemental Fig. 4).

A consensus peak set was obtained for each cell type (Ecto: 55,917 peaks, Endo: 46,535 peaks, VK2: 48,164 peaks). As expected, the majority of open chromatin peaks across all cell types were located in proximal promoter regions (< 1 kb) or intergenic regions (Fig. 4A). To assess cell-specific differences in chromatin organization, we compared normalized sequence read counts across genomic regions. While no differences were observed in quantile-normalized counts between TSS and gene bodies for each cell type (Supplemental Fig. 5), ectocervical epithelial cells uniquely exhibited two distinct clusters of chromatin accessibility in consensus peak regions (Fig. 4B). This clustering was not observed in

endocervical and vaginal epithelial cells (Fig. 4B). Based on these findings and the pronounced epigenetic shifts detected in ectocervical epithelial cells via RNA sequencing, subsequent ATAC-seq analyses focused on ectocervical cells.

Tests for regions of differential accessibility between the treatment conditions were performed in the cell type-specific consensus peak sites (Supplemental Tables 8a-c). Differential accessibility analysis revealed 8,147 regions with altered chromatin accessibility in *L. crispatus* supernatant-treated ectocervical cells, with 8,125 regions showing decreased accessibility and only 22 regions showing increased accessibility. In contrast, endocervical and vaginal cells exhibited far fewer differentially regulated sites (21 and 109 total sites, respectively). In ectocervical cells, regions with decreased accessibility were predominantly located in distal intergenic and intronic regions, with a corresponding reduction in proximal promoter sites (Fig. 4C). Genes neighboring these differentially accessible regions overlapped significantly with those identified as differentially expressed by RNA-sequencing (683/762 genes) [44–46].

To explore the functional relevance of these findings, we assessed whether differentially accessible regions were enriched in tissue-specific regulatory elements or enhancer regions identified by the Encyclopedia of DNA Elements (ENCODE) Project in primary tissues or open accessibility regions identified in different primary cancer specimens [47–49]. Surprisingly, little overlap was observed between differentially accessible regions in ectocervical cells and putative enhancer regions (Fig. 4D, Chi-squared test for trend $p = 0.8334$). This lack of overlap may reflect the absence of primary cervical tissue specimens in the ENCODE dataset, as comparisons were limited to vaginal and uterine samples. In contrast, all accessible chromatin regions in ectocervical epithelial cells showed substantial overlap with published likely enhancer regions (Fig. 4E, Chi-squared test for trend $p = 0.0028$).

Motif analysis of downregulated differentially accessible sites identified enrichment for 497 transcription factor motifs with an FDR of 0.05 (Supplemental Table 9, Fig. 4F). The top five motifs were centrally positioned within the peaks, consistent with potential true transcription factor binding [50]. Gene-disease enrichment analysis of these transcription factors revealed associations with pathways related to neoplasms, endometriosis and infertility, conditions potentially mitigated by *Lactobacillus*-dominated microbiota (Fig. 4G) [51, 52]. Specificity analysis using random transcription factor lists confirmed these findings, as random motifs predominantly enriched for unrelated pathways, such as craniofacial anomalies (Supplemental Fig. 6).

Discussion

Host-microbe interactions are critical determinants of health and disease across multiple biological systems. This study sheds light on unique molecular mechanisms underlying host-microbe interactions within the cervicovaginal space, addressing a significant gap in our understanding of reproductive health. Our findings reveal that *G. vaginalis*, a facultative anaerobic bacteria associated with many gynecological disorders, including STIs [9, 10], cervical cancer [53, 54], infertility [7, 8], and preterm birth

[4, 12], induces diverse immune pathways, dysregulates the innate immune response, and increases epithelial cell death. In contrast, *L. crispatus*, a key species in optimal vaginal microbiomes, promotes epigenetic modifications in ectocervical cells without inducing cell death. Together, these findings highlight the complexity of host-microbe interactions in the lower reproductive tract and reveal distinct molecular pathways by which optimal and non-optimal bacteria contribute to reproductive health and disease.

While some studies have characterized cervicovaginal microbiomes using high-throughput sequencing technologies [33, 55, 56], few have investigated the host transcriptional and functional pathways altered by host-microbe interactions in different cervicovaginal epithelial cell types critical to the function of the lower reproductive tract. Our RNA-seq results demonstrate that *G. vaginalis* and *L. crispatus* culture supernatants modulate distinct host genes and functional pathways in a cell-type-specific manner. These findings reflect the distinct functional diversity of epithelial barrier cells in the lower genital tract. Notably, each cervicovaginal epithelial cell type exhibited unique transcriptomic signatures in response to bacterial culture supernatant exposure, suggesting that specific tissue microenvironments can contribute to the varied reproductive outcomes observed *in vivo.* Understanding the microbial transcriptional activity within these distinct epithelial niches is essential to elucidating microbial functions, and not simply their presence, that drive host responses.

Human studies have shown positive correlations between a pro-inflammatory state and the presence of an anaerobe-rich cervicovaginal microbiome [4, 22, 57–60]. For example, vaginal swabs from Kenyan and Ugandan women with non-optimal microbiomes revealed an upregulation of cytokines involved in the innate immune response [58]. Consistent with these findings, our RNA sequencing results show that exposure to *G. vaginalis* culture supernatant upregulates genes involved in innate immune signaling pathways (cytokines/chemokines), increases anti-microbial peptides (AMPs), and induces cell death in cervicovaginal epithelial cells. While the inflammatory response to *G. vaginalis* has been previously linked to cellular damage and death [22, 32, 61], our study uniquely highlights the role of AMPs in this process. As a critical part of the innate immune response, AMPs, also known as host defense peptides, act to destroy invading pathogens using a variety of biological processes. AMPs, including CCL20, S100A8, SLPI and LCN2, use diverse mechanisms to defend against pathogens. For example, CCL20 promotes immune cell migration [62–64], S100A8 acts as a chemoattractant for neutrophils [65, 66], SLPI protects against neutrophil elastase, and LCN2 sequesters iron to inhibit bacterial growth. While discordant alterations in RNA and protein were noted for some AMPs, the effect of experimental timing on these results can't be ruled out, as RNA and protein levels were assessed at the same time post bacterial exposure. However, notably, CCL20 was upregulated (RNA and protein) across all cell types in response to *G. vaginalis* and may play a critical role in recruiting immune cells to combat bacterial colonization. As a potent chemoattractant of lymphocytes and dendritic cells, CCL20 likely contributes to the recruitment of monocytes to defend against *G. vaginalis* colonization *in vivo.* Interestingly, CCL20 is the only chemokine that interacts with CC chemokine receptor 6 (CCR6), a property shared with the anti-microbial β-defensins. We have previously demonstrated that higher levels of b-defensin-2 were protective against preterm birth in the presence of specific anaerobes common to CST IV [4]. These two

findings suggest that signaling through CCR6 by CCL20 and/or specific AMPs may be critical regulators of the host immune response to non-optimal bacteria. The role of CCL20 and/or other AMPs in limiting *G. vaginalis*-induced cell death and inflammation requires further investigation. Like CCL20, S100A8/A9 (also known as calprotectin) was increased after *G. vaginalis* exposure. In addition to neutrophil recruitment, S100A8/A9 acts to sequester metals/nutrients (calcium, iron, zinc, manganese) [67–69] as part of a process termed nutritional immunity in which metal-chelating host defense mechanisms are used to prevent infection [69, 70]. As most *Lactobacillus* species, except *L. iners*, require manganese for colonization [71, 72], it is possible that *G. vaginalis*-mediated increases in S100A8/A9 could act to limit *Lactobacillus* growth, thus potentially contributing to microbial dysbiosis in the cervicovaginal space. However, very little is known about the biological mechanisms regulating the effects of S100A8/A9 in the lower reproductive tract, and thus, elucidating the role of this AMP in modifying the cervicovaginal microbiome requires additional studies. It is biologically plausible that *G. vaginalis* induces diverse AMPs with opposing effects in the cervicovaginal environment, as evidenced by S100A8/A9's ability to limit the bacterial colonization of optimal bacteria, while CCL20 stimulates the host immune response to reduce non-optimal bacterial colonization. Furthermore, specific AMPs may have multiple distinct functions in the cervicovaginal space. For instance, S100A8/A9 has been shown to promote leukocyte recruitment to initiate a host immune response [73], while simultaneously restricting nutrients essential for the growth of beneficial bacteria such as *Lactobacillus*. Interestingly, *L. crispatus* exposure downregulated the transcription of CCL20 and S100A8 suggesting that AMPs may help to promote *L. crispatus* growth. These findings provide further evidence for an intricate and complex relationship between the host and the vaginal microbiota.

In contrast to *G. vaginalis*, *L. crispatus* is considered an optimal bacterium that promotes reproductive health [2, 74–77]. Providing a plausible biological rationale for this protection, despite possessing a bacterial cell wall that should activate TLR-2, *L. crispatus* does not induce inflammation [22, 32] in part due to the presence of protective S-layer proteins [78]. However, the biological mechanisms underlying the beneficial properties of *L. crispatus* and other vaginal *Lactobacillus* spp. remain largely unknown, limiting the development of therapeutics that could leverage the beneficial properties of this bacteria. This study provides novel insight into these biological mechanisms and suggests that this protection may be mediated through epigenetic modifications. Specifically, *L. crispatus* culture supernatant modulates genes governing transcriptional and epigenetic regulation, leading to global reorganization of the epigenome in cervicovaginal epithelial cells. ATAC-seq analysis showed that exposure to *L. crispatus* culture supernatant reduced the number of open chromatin regions, suggesting a potential mechanism for increasing cervical epithelial cell resilience to infectious agents *(e.g., Chlamydia trachomatis*, HIV, HPV) [79, 80]. Consistent with our findings, a prior study demonstrated that *Lactobacillus* culture supernatants, specifically D-lactic acid, reduced *C. trachomatis* infection by modulating cell proliferation, a process essential for infection, via decreasing histone deacetylase 4 (HDAC4) and increasing histone acetylase EP300 gene expression [55]. While this previous study found that lactic acid induced these epigenetic modifications, further investigation is needed to identify additional metabolites or proteins present in *L. crispatus* culture supernatants that may alter the host epigenome. Since chromatin

accessibility was altered in ectocervical cell lines, additional studies would be needed to determine whether similar chromatin alterations occur *in vivo* or in primary cervical cells. *L. crispatus*-mediated epigenome alterations in cervical epithelial cells supports a role for *L. crispatus* in protecting against microbial pathogens, vaginal infection, and even cervical cancer. [53]

Intriguingly, regions of differential chromatin accessibility did not overlap with known enhancer regions, likely due to the lack of ENCODE or ATAC profiles for primary cervical tissue for comparison. Further, the profiled cervical cancer specimens were derived from only four specimens and likely were a suboptimal reference for comparison to the ectocervical lines used in our study. This suggests that the identified differentially regulated regions are probably cell type-specific enhancers in ectocervical cells. The enrichment of intronic regions among differentially accessible sites provides evidence that *L. crispatus* may regulate isoform transcription through chromatin modulation. While, in this study, RNA sequencing could not address this hypothesis, future work employing long-read RNA sequencing may provide clarity [81]. Additionally, recent research has highlighted lactate, a precursor to acetyl-CoA, as both a precursor for histone acetylation and a direct modifier of histones via lactylation [82–85]. Histone lactylation could lead to increased gene expression putatively associated with more open chromatin. Further exploration of lactate's role in chromatin accessibility and gene regulation could deepen our understanding of these processes. While we are unable to explain how differentially regulated sites contribute to gene regulation, disease gene enrichment analysis of transcription factors associated with the motifs at these sites indicate multiple pathologies related to women's health, including fertility and endometriosis. These results point to an intriguing avenue of study to better understand the molecular underpinnings of these common but poorly understood reproductive disorders.

A limitation of this study is the focus on single vaginal bacterial species and strains. While whole microbiomes are needed to more accurately reflect the cervicovaginal microenvironment, focusing on individual bacteria allowed us to identify specific functional pathways underlying adverse outcomes. Future research should expand to include other high-risk anaerobic bacteria [4], such as *Sneathia*, *Mobiluncus* and *Prevotella* species, to elucidate their combined effects on cervicovaginal functions. Further, these studies will benefit from our foundational findings pointing to specific biological functions of both *G. vaginalis* and *L. crispatus* that could be leveraged to develop mechanistic hypotheses. Additionally, the inclusion of transcriptional profiles, including ATAC-seq data, from normal cervical and vaginal tissues in databases such as ENCODE, is critical for advancing [47] our understanding of host-microbial interactions in the reproductive tract and their role in reproductive outcomes.

In summary, this study identifies novel transcriptional and epigenomic pathways altered by common vaginal bacteria, highlighting the molecular complexity of host-microbial interactions in the cervicovaginal environment and their potential contributions to reproductive health and disease. Using unbiased sequencing approaches, we demonstrated that *G. vaginalis* activates the innate immune response, whereas *L. crispatus* modulates transcription and chromatin accessibility. Additionally, we found that these bacteria alter the transcriptional and epigenomic landscapes in distinct ways across the different epithelial surfaces in the lower reproductive tract. Collectively, our findings highlight

potential therapeutic targets: 1) modulating the inflammatory response associated with *G. vaginalis* to mitigate STIs, bacterial vaginosis and preterm birth, and 2) leveraging *L. crispatus*-mediated epigenetic changes to strengthen cervicovaginal epithelial barriers against viral infections such as HPV. Continued investigations into host-microbial interactions in the female reproductive tract hold great promise for optimizing reproductive health.

Materials and Methods

Cell Culture

Ectocervical (Ect/E6E7, ATCC# CRL-2614) (Ecto), endocervical (End1/E6E7, ATCC# CRL-2615) (Endo) and vaginal (VK2/E6E7, ATCC# CRL-2616) (VK2) human epithelial cell lines (American Type Culture Collection, Manassas, VA) were cultured in Keratinocyte-Serum Free Media (KSFM) supplemented with 0.1 ng/mL epidermal growth factor and 50 µg/mL bovine pituitary extract (Gibco, Life Technologies), 100 U/mL penicillin, and 100 µg/mL of streptomycin at 37$^{\mathrm{o}}$C in a 5% CO_2 humidified incubator.

Bacterial Cultures and Preparation of Bacteria-Free Supernatants

Human clinical isolates of *L. crispatus* (ATCC 33197) or *G. vaginalis* (ATCC 14018), were obtained from the American Type Culture Collection (Manassas, VA). *G. vaginalis* was grown on Tryptic Soy Agar with 5% Sheep Blood plates (Hardy Diagnostics) and *L. crispatus* was grown on De Man, Rogosa and Sharpe agar (Fisher Scientific); both strains were grown in New York City III (NYCIII) broth. Bacteria were grown at 37$^{\mathrm{o}}$C in an anaerobic glove box (Coy Labs, Grass Lake, MI).

For each experiment the following bacterial growth protocol was followed: *L. crispatus* and *G. vaginalis* glycerol stocks were streaked on agar plates, as well as into broth tubes and grown overnight. The broth starter cultures were diluted to an optical density of 0.2 and then used to inoculate 20ml working cultures, which were grown for 20 hours (*G. vaginalis*) to 48 hours (*L. crispatus*) prior to use in experiments. Bacterial densities of the working cultures were estimated the day of the experiment based on optical density readings at 600 nm using an Epoch2 plate reader (Biotek, Winooski, VT), and the appropriate volume was centrifuged at 13,000 x *g* for 3 min. To obtain bacteria-free culture supernatants, the working cultures were centrifuged at 13,000 x *g* for 3 min and the supernatant was filtered through a 0.22 µm filter (Fisher Scientific) to remove any remaining live bacteria. Bacteria-free culture supernatants were diluted to 1% v/v in the appropriate cell culture media without antibiotics.

In vitro Epithelial and Immune Cell - Bacteria Interactions

Ectocervical, endocervical and vaginal cells were plated at 1.5 x 10^5 cells/well in twenty-four well plates containing KSFM without antibiotics. The next day, the cells were exposed to 1% (v/v) *L. crispatus* or *G. vaginalis* bacteria-free culture supernatants (generated from a 1x10^7 CFU/mL culture) for 24 hr. Bacteria-free culture supernatant percentage was based on a dose response (1% vs 10%) (Supplemental Fig. 7).

For cells exposed to 1% bacteria-free culture supernatants from *L. crispatus*, KSFM media was supplemented with 50mM HEPES and sodium bicarbonate (3000 mg/L total concentration) to bring the pH of the media up to a physiological level (7.2). For all supernatant experiments, cells were also exposed to 1% (v/v) NYCIII bacterial growth media alone (diluted in KSFM) to determine any baseline effects of the bacterial growth media on the outcomes of interest. 1% NYCIII (NYC) acted as the control for all bacteria-free culture supernatant exposures. At the end of each experiment, cell culture media was collected for cell death (supplemental methods), ELISA assays and/or the cells were collected in Trizol (Invitrogen, Thermo-Fisher Scientific) for RNA extraction.

RNA Sequencing and Analysis

RNA was extracted from ectocervical, endocervical and vaginal cells after exposure to culture supernatants from *L. crispatus* and *G. vaginalis* (n = 3/treatment group) collected in Trizol using the Qiagen-RNeasy Plus Mini kit by the Penn Next-Generation Sequencing Core. The resulting cell death in these samples is shown in Supplemental Fig. 8. Despite some observed cell death following bacteria supernatant exposure, the resulting RNA had RIN values > 9. Illumina sequencing libraries were prepared using the Illumina TruSeq mRNA stranded library prep kit according to the manufacturer recommendations. The resulting libraries had an average molarity of 69 nM +/1 27 nM. Libraries were sequenced to a median depth of 41 million 100 bp single reads on an Illumina NovaSeq 6000. Transcript quantification from RNA-seq data was performed using Salmon and release 38 (GRCh38.p13) of the human genome [86, 87]. Several Bioconductor packages in R were used for subsequent steps [88, 89]. The output was annotated and summarized using tximeta and further annotation was completed with biomaRt [90, 91]. Principle Component Plots (PCA) were created using pcaExplorer [92, 93]. The normalizations and statistical analyses were done with DESeq2 [94]. Heatmaps for anti-microbial peptides were created using "pheatmap" in R (version 4.1.2). The full RNA-seq dataset was submitted to Gene Expression Omnibus (accession # GSE234837).

RNA-seq Pathway Analysis

PathfindR (v. 1.64) was used for pathway enrichment analysis using Gene Ontology terms (version from 2022-11-03) (https://github.com/egeulgen/pathfindR and https://release.geneontology.org/) [34, 35]. Upregulated and downregulated genes were grouped together for each comparison. The enrichment threshold was set at an FDR of 0.05, with a significant gene threshold of 0.02. A heatmap for enrichment scores for each comparison was created by first calculating and aggregating term scores for each sample included for each comparison and then averaging the scores across all compared samples as previously described [34]. *ComplexHeatmap* package in R (v. 2.14.0) was then used to visualize the comparison of GO term analysis (rows) for all the comparisons (columns) [37]. Rows were clustered by the "complete" method with a kmeans = 5. A word cloud was used to represent the most significant recurring pathways in a cluster. Generic terms or single letters were excluded from word cloud ("pathway", "cellular", "regulation", "positive", "negative", "cell", "complex", "process", "factor", "activity",

"protein", "dna", "rna", "levels", "binding", "response", "signaling", "receptor", "production", "t", "ii", "p", "g", "c", "via", "class").

ELISA

Ectocervical, endocervical or vaginal, cells were cultured in 24-well plates and exposed to bacterial culture supernatants as stated above. Anti-microbial peptides, CCL20, SLPI, LCN2, S100A8/A9, were measured in cell culture media after 24 hours of exposure (n = 3/group with n = 3 technical replicates per experiment). The expression of these analytes was measured by a ligand-specific commercially available ELISA kit that utilizes a quantitative sandwich enzyme immunoassay technique using reagents from R&D Systems (Minneapolis, MN).

ATAC-seq Nuclei Extraction, Tagmentation, Purification and Library Amplification

ATAC-seq was performed on ectocervical, endocervical and vaginal cells after exposure to *L. crispatus* bacteria-free supernatants (n = 3/treatment group). ATAC-seq libraries were generated using the ATAC-seq Kit from Diagenode (Diagenode, A Hologic Company) according to manufacturer instructions. Briefly, nuclei were extracted from 50,000 cells. Tagmentation was completed by resuspending the isolated nuclei in transposase reaction mix and the samples were purified using the kit's provided columns. Following purification, library fragments were amplified by PCR according to the manufacturer recommendations. Unique Dual Indexes Primer Pairs were incorporated for multiplexed sequencing. To reduce amplification bias, after the first 5 cycles of the PCR reaction, qPCR was used to determine how many additional cycles were needed to produce enough library to meet the required amount for sequencing. For this, an aliquot of the PCR reaction was added to Sybr Green and amplified for 20 cycles. Libraries were amplified for a total of 11–13 cycles (with one library requiring 17 cycles for amplification). Final libraries were purified using bead purification (Beckman Coulter), then assessed for size distribution and concentration using a BioAnalyzer High Sensitivity DNA Kit (Agilent Technologies). The resulting libraries were pooled. The pool was diluted to 2 nM, denatured, and the 13 libraries were loaded onto an S1-100 (2x50) flow cell on an Illumina NovaSeq 6000 (Illumina, Inc.) according to the manufacturer's instructions. The average read number per sample was 50M+/- 20%. De-multiplexed and adapter-trimmed sequencing reads were generated using bcl2fastq. The full ATAC-seq dataset was submitted to Gene Expression Omnibus (accession # GSE233444).

ATAC-seq Mapping and Peak Calling

ATAC-seq data analysis was adapted from a previously published approach using PEPATAC (v. 0.10.3) [41]. Peaks for each cervicovaginal line were called separately to allow for cell type specific differences in chromatin accessibility pattern. In brief, raw FASTQ files were processed and mapped to release 38 (GRCh38.p13) of the human genome using the PEPATAC pipeline [95]. Reads were trimmed with *skewer*

and then aligned with *bowtie2* using default settings [96, 97]. Duplicate reads were removed using *samblaster* [98].

An iterative overlap peak calling strategy on fixed-sized peaks of 501 bp was used to define a set number of peaks for each cell type for downstream differential accessibility comparison [41]. First, for each biological replicate, MACS2 was used to call peaks with the parameters as follows: --peak-type fixed --extend 250 [99]. Biological replicates of each treatment and then both treatments together from each cell type were merged using an iterative overlap approach previously described [41, 49]. Blacklisted regions were excluded from called peaks (accessed 4 November 2022 at https://github.com/Boyle-Lab/Blacklist) [100].

ATAC-seq Peak and Differential Accessibility Analysis

Peak location was annotated with *CHIPseeker* (v 1.30.3) [101]. Counts for peaks were calculated using *Rsubread* (v. 2.8.2)[102]. We determined the differential accessibility of peaks between treatments with *DESeq2* (v. 1.34.0)[103]. We compared *L. crispatus* culture supernatant treated to NYCIII media controls for each cell type. A Wald test was used to determine significance. A peak was defined as statistically significant in differential accessibility if |log2foldchange| > 1 and FDR < 0.05. We utilized the R package *rGREAT* (v. 1.99.0) for the nearest gene analysis to access the Genome Regions Enrichment of Annotations Tool (GREAT) web service [44–46]. For GREAT, we used the parameters for "the two closest genes" to a differential accessible site as it is frequently not the closest genes that is differentially regulated.

Motif analysis was performed using Simple Enrichment Analysis version 5.5 as part of the MEME Suite (https://meme-suite.org/meme/tools/sea) [104, 105]. Differentially accessible sites were inputted, and the CIS-BP 2.0 motifs database were used for the query [106]. Gene-disease enrichment analysis was performed using *disgenet2R* (v. 0.99.2, https://www.disgenet.org/) [107]. Random gene lists were generated for comparison by sampling 497 transcription factors from the CIS-BP 2.0 database to ascertain the baseline disease enrichment bias of the database.

ATAC-seq Chromatin Accessibility Visualization

EasSeq (v1) was utilized to visualize the data [108]. Biological replicates of BAM files were pooled for quantification of specific regions. Quantile normalization was used for counts per region for visualization to minimize bias from sequencing depth. Calculation of overlap performed both by any amount of overlap and the exact overlap of base pairs between all comparisons. Random regions for comparisons to differentially accessible regions or all ectocervical open chromatin regions were generated by Regulatory Sequence Analysis Tools matched for each cell type by number of fragments, fragment size, and GC content (random genome fragments tool; http://rsat.sb-roscoff.fr/) [109].

ENCODE datasets for all human enhancer-like sequences (ELS, defined as high DNAse-seq signal and high H3K37me3), or tissue-specific regulators were obtained from https://screen.encodeproject.org/ [47, 48]. For uterus and vaginal specimens, "Low-DNase" were filtered out to enrich for sites that had any

evidence of potential enhancer or regulator activity. However, strict enhancer-like signature criteria could not be applied because all sequencing modalities were not available for all the samples. Primary cancer cell data sets were obtained from supplemental of published ATAC profiling [49]. Chi square analysis of number of overlapping sites was performed by Graph Pad.

Statistical Analysis

Statistical analyses were performed for all experiments (except for RNA or ATAC sequencing, statistical analysis is described above for each) with the GraphPad Prism Software (Version 9.0, San Diego, CA). For data that were normally distributed (as assessed by Shapiro-Wilk test), one-way analysis of variance (ANOVA) was performed. If statistical significance was reached ($p < 0.05$), then pair-wise comparison with a Tukey post hoc test was performed for multiple comparisons. If data were not normally distributed, then the non-parametric Kruskal-Wallis test was used and pairwise comparison was done using Dunn's multiple comparison test. Chi test for trend was utilized to compare overlaps of indicated ectocervical peaks with the number of a random set of sites matched for size and CG content.

Declarations

Ethics approval and consent to participate

No animal or human samples were used in this study therefore this is not applicable for this manuscript.

Consent for publication

Not applicable for this manuscript.

Funding

Availability of data and materials

The full RNA-seq dataset was submitted to Gene Expression Omnibus (accession # GSE234837). The full ATAC-seq dataset was submitted to Gene Expression Omnibus (accession # GSE233444). All other data is provided within the manuscript or supplementary information files.

Competing interests

MAE is a consultant with equity for MIRVIE. JR is co-founder of LUCA Biologics, a biotechnology company focusing on translating microbiome research into live biotherapeutics drugs for women's

health. JR is a Scientific Advisor for Ancilia Bio. All other authors declare that they have no competing interests.

Author contributions

LA, AGC, JR and MAE wrote the manuscript and created the figures. LA, AGC and MAE conceived and designed the experiments. LA and BF performed the experiments. LA, AGC and MAE analyzed the data. LA, AGC, JR and MAE contributed scientific ideas and project direction. All authors contributed to manuscript revision and approved the final manuscript.

Acknowledgements

References

1. Nakama C, Thompson B, Szybala C, McBeth A, Dobner P, Zwickey H. The continuum of microbial ecosystems along the female reproductive tract: Implications for health and fertility. Pathogens. 2022; 11:1244.
2. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, et al. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci U S A. 2011; 108 Suppl 1:4680–4687.
3. Consortium. HMP. Structure, function and diversity of the healthy human microbiome. Nature. 2012; 486:207–214.
4. Elovitz MA, Gajer P, Riis V, Brown AG, Humphrys MS, Holm JB, Ravel J. Cervicovaginal microbiota and local immune response modulate the risk of spontaneous preterm delivery. Nat Commun. 2019; 10:1305–1305.
5. France MT, Ma B, Gajer P, Brown S, Humphrys MS, Holm JB, Waetjen LE, Brotman RM, Ravel J. Valencia: A nearest centroid classification method for vaginal microbial communities based on composition. Microbiome. 2020; 8:166.
6. Gajer P, Brotman RM, Bai G, Sakamoto J, Schütte UME, Zhong X, Koenig SSK, Fu L, Ma Z, Zhou X, et al. Temporal dynamics of the human vaginal microbiota. Science Translational Medicine. 2012; 4:132ra152-132ra152.
7. Campisciano G, Florian F, D'Eustacchio A, Stanković D, Ricci G, De Seta F, Comar M. Subclinical alteration of the cervical–vaginal microbiome in women with idiopathic infertility. Journal of Cellular Physiology. 2017; 232:1681–1688.

8. van Oostrum N, De Sutter P, Meys J, Verstraelen H. Risks associated with bacterial vaginosis in infertility patients: A systematic review and meta-analysis. Human Reproduction. 2013; 28:1809–1815.

9. Brusselaers N, Shrestha S, van de Wijgert J, Verstraelen H. Vaginal dysbiosis and the risk of human papillomavirus and cervical cancer: Systematic review and meta-analysis. Am J Obstet Gynecol. 2019; 221:9–18.e18.

10. Price JT, Vwalika B, Hobbs M, Nelson JAE, Stringer EM, Zou F, Rittenhouse KJ, Azcarate-Peril A, Kasaro MP, Stringer JSA. Highly diverse anaerobe-predominant vaginal microbiota among hiv-infected pregnant women in zambia. PLoS One. 2019; 14:e0223128.

11. Elovitz MA, Gajer P, Riis V, Brown AG, Humphrys MS, Holm JB, Ravel J. Cervicovaginal microbiota and local immune response modulate the risk of spontaneous preterm delivery. Nat Commun. 2019; 10:1305.

12. Callahan BJ, DiGiulio DB, Goltsman DSA, Sun CL, Costello EK, Jeganathan P, Biggio JR, Wong RJ, Druzin ML, Shaw GM, et al. Replication and refinement of a vaginal microbial signature of preterm birth in two racially distinct cohorts of us women. Proc Natl Acad Sci U S A. 2017; 114:9966–9971.

13. Potter RF, Burnham CD, Dantas G. In silico analysis of gardnerella genomospecies detected in the setting of bacterial vaginosis. Clin Chem. 2019; 65:1375–1387.

14. Forsberg J-G. Cervicovaginal epithelium: Its origin and development. American Journal of Obstetrics and Gynecology. 1973; 115:1025–1043.

15. Fritsch H, Hoermann R, Bitsche M, Pechriggl E, Reich O. Development of epithelial and mesenchymal regionalization of the human fetal utero-vaginal anlagen. Journal of Anatomy. 2013; 222:462–472.

16. Kurita T. Normal and abnormal epithelial differentiation in the female reproductive tract. Differentiation. 2011; 82:117–126.

17. Taherali F, Varum F, Basit AW. A slippery slope: On the origin, role and physiology of mucus. Advanced Drug Delivery Reviews. 2018; 124:16–33.

18. Nazli A, Chan O, Dobson-Belaire WN, Ouellet M, Tremblay MJ, Gray-Owen SD, Arsenault AL, Kaushic C. Exposure to hiv-1 directly impairs mucosal epithelial barrier integrity allowing microbial translocation. PLoS Pathog. 2010; 6:e1000852.

19. Cherne MD, Cole AL, Newberry L, Schmidt-Owens M, Deichen M, Cole AM. Matrix metalloproteinases expressed in response to bacterial vaginosis disrupt the endocervical epithelium, increasing transmigration of hiv. Infect Immun. 2020; 88.

20. Nazli A, Kafka JK, Ferreira VH, Anipindi V, Mueller K, Osborne BJ, Dizzell S, Chauvin S, Mian MF, Ouellet M, et al. Hiv-1 gp120 induces tlr2- and tlr4-mediated innate immune activation in human female genital epithelium. J Immunol. 2013; 191:4246–4258.

21. Benjelloun F, Quillay H, Cannou C, Marlin R, Madec Y, Fernandez H, Chrétien F, Le Grand R, Barré-Sinoussi F, Nugeyre M-T, Menu E. Activation of toll-like receptors differentially modulates inflammation in the human reproductive tract: Preliminary findings. Frontiers in Immunology. 2020; 11.

22. Anton L, Ferguson B, Friedman ES, Gerson KD, Brown AG, Elovitz MA. Gardnerella vaginalis alters cervicovaginal epithelial cell function through microbe-specific immune responses. Microbiome. 2022; 10:119.

23. Zariffard MR, Novak RM, Lurain N, Sha BE, Graham P, Spear GT. Induction of tumor necrosis factor-alpha secretion and toll-like receptor 2 and 4 mrna expression by genital mucosal fluids from women with bacterial vaginosis. J Infect Dis. 2005; 191:1913–1921.

24. O'Hanlon DE, Moench TR, Cone RA. Vaginal ph and microbicidal lactic acid when lactobacilli dominate the microbiota. PLoS One. 2013; 8:e80074.

25. Alakomi HL, Skyttä E, Saarela M, Mattila-Sandholm T, Latva-Kala K, Helander IM. Lactic acid permeabilizes gram-negative bacteria by disrupting the outer membrane. Appl Environ Microbiol. 2000; 66:2001–2005.

26. O'Hanlon DE, Moench TR, Cone RA. In vaginal fluid, bacteria associated with bacterial vaginosis can be suppressed with lactic acid but not hydrogen peroxide. BMC Infect Dis. 2011; 11:200.

27. Tyssen D, Wang YY, Hayward JA, Agius PA, DeLong K, Aldunate M, Ravel J, Moench TR, Cone RA, Tachedjian G. Anti-hiv-1 activity of lactic acid in human cervicovaginal fluid. mSphere. 2018; 3.

28. Aldunate M, Tyssen D, Johnson A, Zakir T, Sonza S, Moench T, Cone R, Tachedjian G. Vaginal concentrations of lactic acid potently inactivate hiv. J Antimicrob Chemother. 2013; 68:2015–2025.

29. Aldunate M, Srbinovski D, Hearps AC, Latham CF, Ramsland PA, Gugasyan R, Cone RA, Tachedjian G. Antimicrobial and immune modulatory effects of lactic acid and short chain fatty acids produced by vaginal microbiota associated with eubiosis and bacterial vaginosis. Frontiers in physiology. 2015; 6:164.

30. Rose WA, 2nd, McGowin CL, Spagnuolo RA, Eaves-Pyles TD, Popov VL, Pyles RB. Commensal bacteria modulate innate immune responses of vaginal epithelial cell multilayer cultures. PLoS One. 2012; 7:e32728.

31. Hearps AC, Tyssen D, Srbinovski D, Bayigga L, Diaz DJD, Aldunate M, Cone RA, Gugasyan R, Anderson DJ, Tachedjian G. Vaginal lactic acid elicits an anti-inflammatory response from human cervicovaginal epithelial cells and inhibits production of pro-inflammatory mediators associated with hiv acquisition. Mucosal Immunology. 2017; 10:1480–1490.

32. Anton L, Sierra LJ, DeVine A, Barila G, Heiser L, Brown AG, Elovitz MA. Common cervicovaginal microbial supernatants alter cervical epithelial function: Mechanisms by which lactobacillus crispatus contributes to cervical health. Front Microbiol. 2018; 9:2181.

33. Delgado-Diaz DJ, Jesaveluk B, Hayward JA, Tyssen D, Alisoltani A, Potgieter M, Bell L, Ross E, Iranzadeh A, Allali I, et al. Lactic acid from vaginal microbiota enhances cervicovaginal epithelial barrier integrity by promoting tight junction protein expression. Microbiome. 2022; 10:141.

34. Ulgen E, Ozisik O, Sezerman OU. Pathfindr: An r package for comprehensive identification of enriched pathways in omics data through active subnetworks. Front Genet. 2019; 10:858.

35. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: Tool for the unification of biology. The gene ontology consortium. Nat

Genet. 2000; 25:25–29.

36. Gu Z. Complex heatmap visualization. iMeta. 2022; 1.

37. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016; 32:2847–2849.

38. Mahajan G, Doherty E, To T, Sutherland A, Grant J, Junaid A, Gulati A, LoGrande N, Izadifar Z, Timilsina SS, et al. Vaginal microbiome-host interactions modeled in a human vagina-on-a-chip. Microbiome. 2022; 10:201.

39. Matsuda E, Takada K, Kobayashi O, Nakajima T, Ikeda Y, Asai-Sato M, Kawakami K, Komatsu A, Chishima F, Komine-Aizawa S, et al. Pathological roles of antimicrobial peptides and pro-inflammatory factors secreted from the cervical epithelium in gardnerella vaginalis-abundant vaginal flora in pregnancy. J Reprod Immunol. 2023; 156:103797.

40. Sierra LJ, Brown AG, Barilá GO, Anton L, Barnum CE, Shetye SS, Soslowsky LJ, Elovitz MA. Colonization of the cervicovaginal space with gardnerella vaginalis leads to local inflammation and cervical remodeling in pregnant mice. PLoS One. 2018; 13:e0191524.

41. Grandi FC, Modi H, Kampman L, Corces MR. Chromatin accessibility profiling by atac-seq. Nat Protoc. 2022; 17:1518–1552.

42. Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Cesari F, Lee C, Almouzni G, Schneider R, Surani MA. Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature. 2008; 452:877–881.

43. Yan F, Powell DR, Curtis DJ, Wong NC. From reads to insight: A hitchhiker's guide to atac-seq data analysis. Genome Biol. 2020; 21:22.

44. Gu Z, Hubschmann D. Rgreat: An r/bioconductor package for functional enrichment on genomic regions. Bioinformatics. 2023; 39.

45. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. Great improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010; 28:495–501.

46. Tanigawa Y, Dyer ES, Bejerano G. Whichtf is functionally important in your open chromatin data? PLoS Comput Biol. 2022; 18:e1010378.

47. Luo Y, Hitz BC, Gabdank I, Hilton JA, Kagda MS, Lam B, Myers Z, Sud P, Jou J, Lin K, et al. New developments on the encyclopedia of DNA elements (encode) data portal. Nucleic Acids Res. 2020; 48:D882-D889.

48. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74.

49. Corces MR, Granja JM, Shams S, Louie BH, Seoane JA, Zhou W, Silva TC, Groeneveld C, Wong CK, Cho SW, et al. The chromatin accessibility landscape of primary human cancers. Science. 2018; 362.

50. Sung MH, Guertin MJ, Baek S, Hager GL. Dnase footprint signatures are dictated by factor dynamics and DNA sequence. Mol Cell. 2014; 56:275–285.

51. Khodaverdi S, Mohammadbeigi R, Khaledi M, Mesdaghinia L, Sharifzadeh F, Nasiripour S, Gorginzadeh M. Beneficial effects of oral lactobacillus on pain severity in women suffering from endometriosis: A pilot placebo-controlled randomized clinical trial. Int J Fertil Steril. 2019; 13:178–183.

52. Sobstyl M, Brecht P, Sobstyl A, Mertowska P, Grywalska E. The role of microbiota in the immunopathogenesis of endometrial cancer. Int J Mol Sci. 2022; 23.

53. Ma Y, Li Y, Liu Y, Cao L, Han X, Gao S, Zhang C. Vaginal microbiome dysbiosis is associated with the different cervical disease status. J Microbiol. 2023; 61:423–432.

54. Mikamo H, Sato Y, Hayasaki Y, Kawazoe K, Izumi K, Ito K, Yamada Y, Tamaya T. Intravaginal bacterial flora in patients with uterine cervical cancer. High incidence of detection of gardnerella vaginalis. J Infect Chemother. 1999; 5:82–85.

55. Edwards VL, Smith SB, McComb EJ, Tamarelle J, Ma B, Humphrys MS, Gajer P, Gwilliam K, Schaefer AM, Lai SK, et al. The cervicovaginal microbiota-host interaction modulates chlamydia trachomatis infection. mBio. 2019; 10.

56. Castro J, Machado D, Cerca N. Unveiling the role of gardnerella vaginalis in polymicrobial bacterial vaginosis biofilms: The impact of other vaginal pathogens living as neighbors. ISME J. 2019; 13:1306–1317.

57. Anahtar Melis N, Byrne Elizabeth H, Doherty Kathleen E, Bowman Brittany A, Yamamoto Hidemi S, Soumillon M, Padavattan N, Ismail N, Moodley A, Sabatini Mary E, et al. Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. Immunity. 2015; 42:965–976.

58. Berard AR, Brubaker DK, Birse K, Lamont A, Mackelprang RD, Noël-Romas L, Perner M, Hou X, Irungu E, Mugo N, et al. Vaginal epithelial dysfunction is mediated by the microbiome, metabolome, and mtor signaling. Cell Reports. 2023; 42:112474.

59. Chan D, Bennett PR, Lee YS, Kundu S, Teoh TG, Adan M, Ahmed S, Brown RG, David AL, Lewis HV, et al. Microbial-driven preterm labour involves crosstalk between the innate and adaptive immune response. Nat Commun. 2022; 13:975.

60. Mitchell C, Marrazzo J. Bacterial vaginosis and the cervicovaginal immune response. Am J Reprod Immunol. 2014; 71:555–563.

61. Vick EJ, Park HS, Huff KA, Brooks KM, Farone AL, Farone MB. Gardnerella vaginalis triggers nlrp3 inflammasome recruitment in thp-1 monocytes. J Reprod Immunol. 2014; 106:67–75.

62. Zaballos A, Varona R, Gutiérrez J, Lind P, Márquez G. Molecular cloning and rna expression of two new human chemokine receptor-like genes. Biochem Biophys Res Commun. 1996; 227:846–853.

63. Hieshima K, Imai T, Opdenakker G, Van Damme J, Kusuda J, Tei H, Sakaki Y, Takatsuki K, Miura R, Yoshie O, Nomiyama H. Molecular cloning of a novel human cc chemokine liver and activation-regulated chemokine (larc) expressed in liver. Chemotactic activity for lymphocytes and gene localization on chromosome 2. J Biol Chem. 1997; 272:5846–5853.

64. Liao F, Alderson R, Su J, Ullrich SJ, Kreider BL, Farber JM. Strl22 is a receptor for the cc chemokine mip-3α. Biochemical and Biophysical Research Communications. 1997; 236:212–217.

65. Yano J, Palmer GE, Eberle KE, Peters BM, Vogl T, McKenzie AN, Fidel PL. Vaginal epithelial cell-derived s100 alarmins induced by candida albicans via pattern recognition receptor interactions are sufficient but not necessary for the acute neutrophil response during experimental vaginal candidiasis. Infection and Immunity. 2014; 82:783–792.

66. Yano J, Lilly E, Barousse M, Fidel PL, Jr. Epithelial cell-derived s100 calcium-binding proteins as key mediators in the hallmark acute neutrophil response during candida vaginitis. Infect Immun. 2010; 78:5126–5137.

67. Corbin BD, Seeley EH, Raab A, Feldmann J, Miller MR, Torres VJ, Anderson KL, Dattilo BM, Dunman PM, Gerads R, et al. Metal chelation and inhibition of bacterial growth in tissue abscesses. Science. 2008; 319:962–965.

68. Damo SM, Kehl-Fie TE, Sugitani N, Holt ME, Rathi S, Murphy WJ, Zhang Y, Betz C, Hench L, Fritz G, et al. Molecular basis for manganese sequestration by calprotectin and roles in the innate immune response to invading bacterial pathogens. Proc Natl Acad Sci U S A. 2013; 110:3841–3846.

69. Hood MI, Skaar EP. Nutritional immunity: Transition metals at the pathogen–host interface. Nature Reviews Microbiology. 2012; 10:525–537.

70. Weinberg ED. Nutritional immunity: Host's attempt to withhold iron from microbial invaders. JAMA. 1975; 231:39–41.

71. Archibald F. Manganese: Its acquisition by and function in the lactic acid bacteria. Crit Rev Microbiol. 1986; 13:63–109.

72. Huynh U, Qiao M, King J, Trinh B, Valdez J, Haq M, Zastrow ML. Differential effects of transition metals on growth and metal uptake for two distinct lactobacillus species. Microbiol Spectr. 2022; 10:e0100621.

73. Wang S, Song R, Wang Z, Jing Z, Wang S, Ma J. S100a8/a9 in inflammation. Front Immunol. 2018; 9:1298.

74. Wan B, Wei LJ, Tan TM, Qin L, Wang H. Inhibitory effect and mechanism of lactobacillus crispatus on cervical precancerous cells ect1/e6e7 and screening of early warning factors. Infect Agent Cancer. 2023; 18:5.

75. Fettweis JM, Serrano MG, Brooks JP, Edwards DJ, Girerd PH, Parikh HI, Huang B, Arodz TJ, Edupuganti L, Glascock AL, et al. The vaginal microbiome and preterm birth. Nat Med. 2019; 25:1012–1021.

76. Anahtar MN, Gootenberg DB, Mitchell CM, Kwon DS. Cervicovaginal microbiota and reproductive health: The virtue of simplicity. Cell Host & Microbe. 2018; 23:159–168.

77. Tachedjian G, Aldunate M, Bradshaw CS, Cone RA. The role of lactic acid production by probiotic lactobacillus species in vaginal health. Research in Microbiology. 2017; 168:782–792.

78. Decout A, Krasias I, Roberts L, Gimeno Molina B, Charenton C, Brown Romero D, Tee QY, Marchesi JR, Ng S, Sykes L, et al. Lactobacillus crispatus s-layer proteins modulate innate immune response

and inflammation in the lower female reproductive tract. Nat Commun. 2024; 15:10879.

79. Crimi E, Benincasa G, Figueroa-Marrero N, Galdiero M, Napoli C. Epigenetic susceptibility to severe respiratory viral infections and its therapeutic implications: A narrative review. Br J Anaesth. 2020; 125:1002–1017.

80. Vicente-Duenas C, Hauer J, Cobaleda C, Borkhardt A, Sanchez-Garcia I. Epigenetic priming in cancer initiation. Trends Cancer. 2018; 4:408–417.

81. Stark R, Grzelak M, Hadfield J. Rna sequencing: The teenage years. Nat Rev Genet. 2019; 20:631–656.

82. Torrini C, Nguyen TTT, Shu C, Mela A, Humala N, Mahajan A, Seeley EH, Zhang G, Westhoff MA, Karpel-Massler G, et al. Lactate is an epigenetic metabolite that drives survival in model systems of glioblastoma. Mol Cell. 2022; 82:3061–3076 e3066.

83. Dai X, Lv X, Thompson EW, Ostrikov KK. Histone lactylation: Epigenetic mark of glycolytic switch. Trends Genet. 2022; 38:124–127.

84. Hagihara H, Shoji H, Otabi H, Toyoda A, Katoh K, Namihira M, Miyakawa T. Protein lactylation induced by neural excitation. Cell Rep. 2021; 37:109820.

85. Zhang D, Tang Z, Huang H, Zhou G, Cui C, Weng Y, Liu W, Kim S, Lee S, Perez-Neut M, et al. Metabolic regulation of gene expression by histone lactylation. Nature. 2019; 574:575–580.

86. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017; 14:417–419.

87. Frankish A, Diekhans M, Jungreis I, Lagarde J, Loveland JE, Mudge JM, Sisu C, Wright JC, Armstrong J, Barnes I, et al. Gencode 2021. Nucleic Acids Res. 2021; 49:D916-d923.

88. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, et al. Bioconductor: Open software development for computational biology and bioinformatics. Genome Biology. 2004; 5:R80.

89. Team RC. R: A language and environment for statistical computing. R foundation for statistical computing. 2021.

90. Love MI, Soneson C, Hickey PF, Johnson LK, Pierce NT, Shepherd L, Morgan M, Patro R. Tximeta: Reference sequence checksums for provenance identification in rna-seq. PLOS Computational Biology. 2020; 16:e1007664.

91. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the r/bioconductor package biomart. Nat Protoc. 2009; 4:1184–1191.

92. Marini F, Binder H. Pcaexplorer: An r/bioconductor package for interacting with rna-seq principal components. BMC Bioinformatics. 2019; 20:331.

93. Ludt A, Ustjanzew A, Binder H, Strauch K, Marini F. Interactive and reproducible workflows for exploring and modeling rna-seq data with pcaexplorer, ideal, and genetonic. Current Protocols. 2022; 2:e411.

94. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. Genome Biology. 2014; 15:550.

95. Smith JP, Corces MR, Xu J, Reuter VP, Chang HY, Sheffield NC. Pepatac: An optimized pipeline for atac-seq data analysis with serial alignments. NAR Genom Bioinform. 2021; 3:lqab101.

96. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. Nat Methods. 2012; 9:357–359.

97. Jiang H, Lei R, Ding SW, Zhu S. Skewer: A fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics. 2014; 15:182.

98. Faust GG, Hall IM. Samblaster: Fast duplicate marking and structural variant read extraction. Bioinformatics. 2014; 30:2503–2505.

99. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. Model-based analysis of chip-seq (macs). Genome Biol. 2008; 9:R137.

100. Amemiya HM, Kundaje A, Boyle AP. The encode blacklist: Identification of problematic regions of the genome. Sci Rep. 2019; 9:9354.

101. Yu G, Wang LG, He QY. Chipseeker: An r/bioconductor package for chip peak annotation, comparison and visualization. Bioinformatics. 2015; 31:2382–2383.

102. Liao Y, Smyth GK, Shi W. The r package rsubread is easier, faster, cheaper and better for alignment and quantification of rna sequencing reads. Nucleic Acids Res. 2019; 47:e47.

103. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. Genome Biol. 2014; 15:550.

104. Bailey TL, Grant CE. Sea: Simple enrichment analysis of motifs. bioRxiv. 2021:2021.2008.2023.457422.

105. Bailey TL, Johnson J, Grant CE, Noble WS. The meme suite. Nucleic Acids Res. 2015; 43:W39-49.

106. Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, Najafabadi HS, Lambert SA, Mann I, Cook K, et al. Determination and inference of eukaryotic transcription factor sequence specificity. Cell. 2014; 158:1431–1443.

107. Pinero J, Ramirez-Anguita JM, Sauch-Pitarch J, Ronzano F, Centeno E, Sanz F, Furlong LI. The disgenet knowledge platform for disease genomics: 2019 update. Nucleic Acids Res. 2020; 48:D845-D855.

108. Lerdrup M, Johansen JV, Agrawal-Singh S, Hansen K. An interactive environment for agile analysis and visualization of chip-sequencing data. Nat Struct Mol Biol. 2016; 23:349–357.

109. Nguyen NTT, Contreras-Moreira B, Castro-Mondragon JA, Santana-Garcia W, Ossio R, Robles-Espinoza CD, Bahin M, Collombet S, Vincens P, Thieffry D, et al. Rsat 2018: Regulatory sequence analysis tools 20th anniversary. Nucleic Acids Res. 2018; 46:W209-W214.
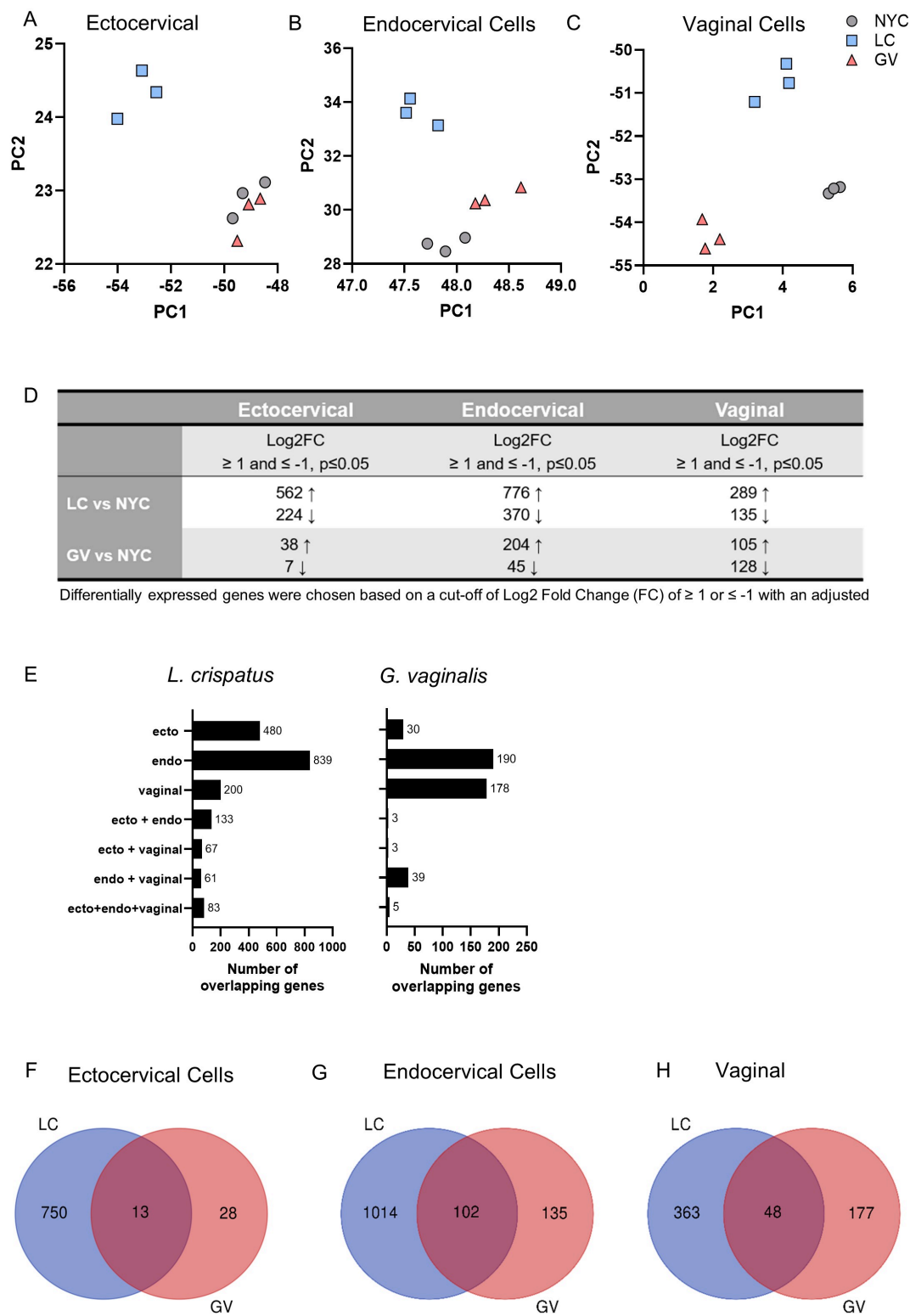
Figures

Figure 1

RNA-seq identified differentially expressed genes in cervicovaginal epithelial cells after 24 hr exposure to culture supernatants from *L. crispatus* or *G. vaginalis*. Principle component analysis (PCA) plots showing gene expression modulation in ectocervical (A), endocervical (B) and vaginal (C) cells exposed to either *G. vaginalis* or *L. crispatus* culture supernatants (vs NYCIII control). The total number of differentially expressed genes (adj. p<0.05, Log2FoldChange ≥ 1 and ≤ -1) in each exposure group by cell

line (D). The number of overlapping differentially expressed genes between cervicovaginal cell types for each bacterial exposure (E-J). The number of overlapping differentially expressed genes between bacterial exposures within each cervicovaginal cell types (F-H).
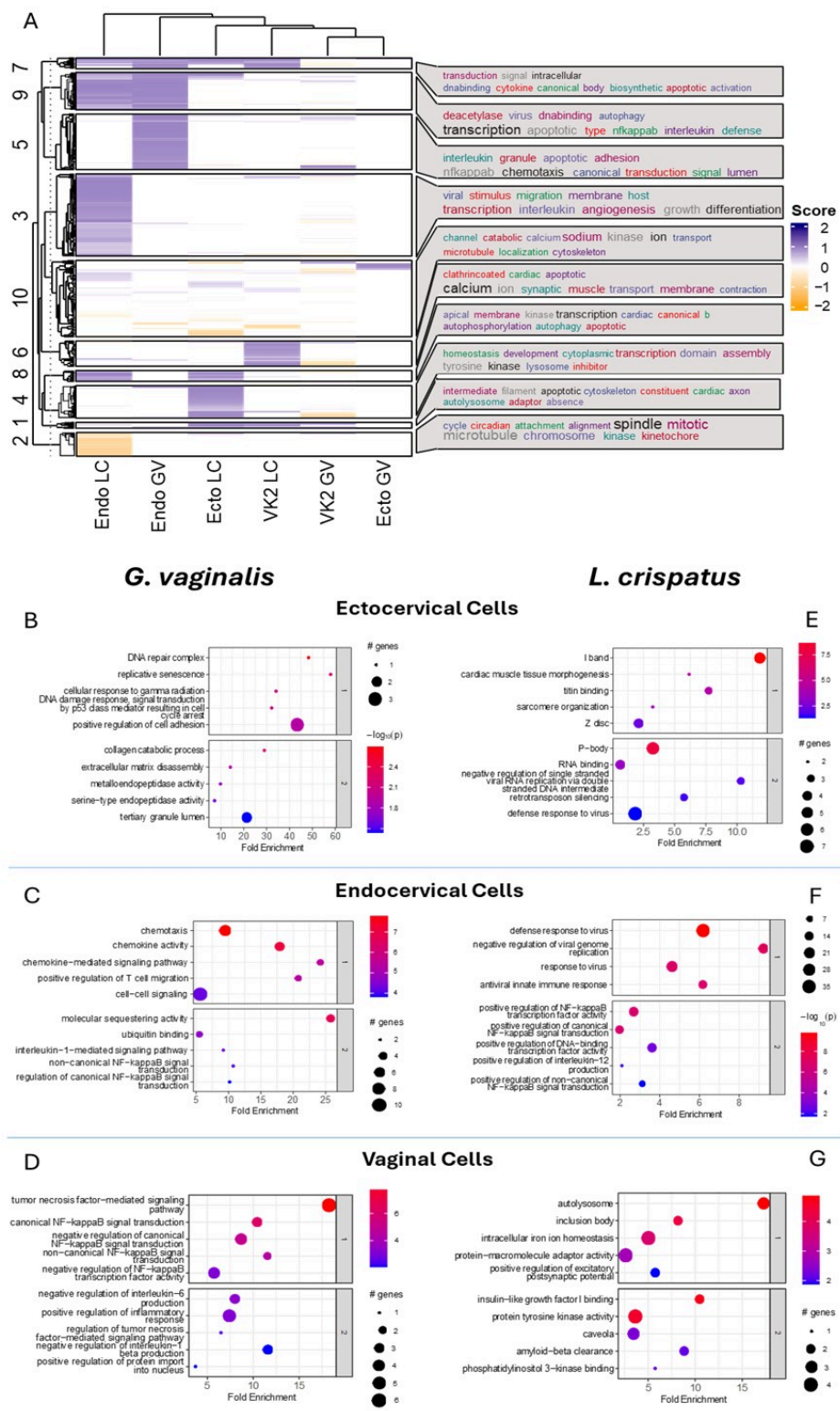

Figure 2

Differential clustering of significant differentially expressed genes (adj. p<0.05, Log2FoldChange ≥ 1 and ≤ -1) between exposure groups and across cervicovaginal cell lines reveal modulation of functional

pathways (A). Functional pathway analysis (B-G) of RNA-seq data for ectocervical (B, E), endocervical (C, F) and vaginal (D, G) epithelial cells exposed to *G. vaginalis* or *L. crispatus*culture supernatants.
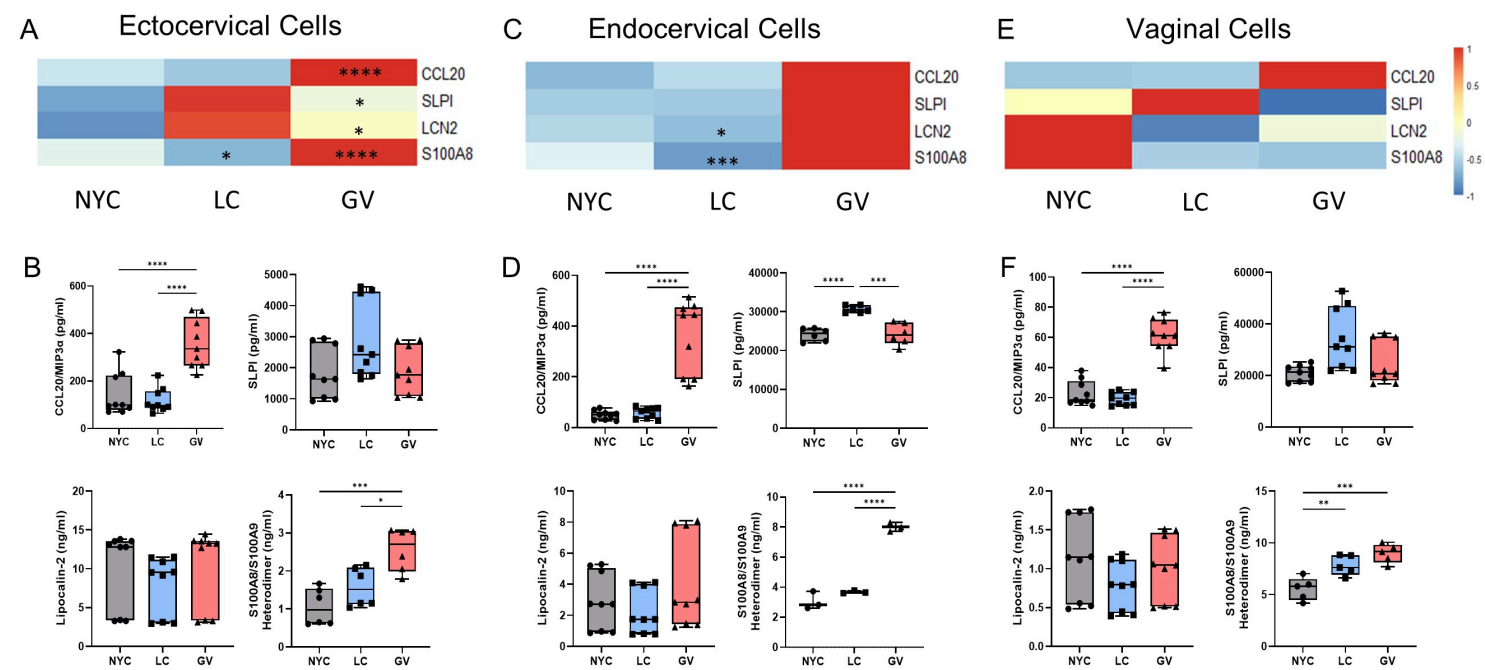

Figure 3

Anti-microbial peptide gene expression and proteins are significantly increased after cervicovaginal cell exposure to *G. vaginalis* culture supernatants. RNA-sequencing identified CCL20, SLPI, LCN2 and S100A8 as being significantly upregulated by *G. vaginalis* culture supernatants in ectocervical (A), endocervical (C) and vaginal (E) epithelial cells. Heatmaps represent differences in normalized reads across rows with significance denoted by asterisks. ELISAs further identified that CCL20 and S100A8 protein levels were increased after *G. vaginalis* culture supernatant exposure (B, D, F). Values are mean ± SEM. Asterisks over solid lines represent comparisons between control (NYC) and treatment groups. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.
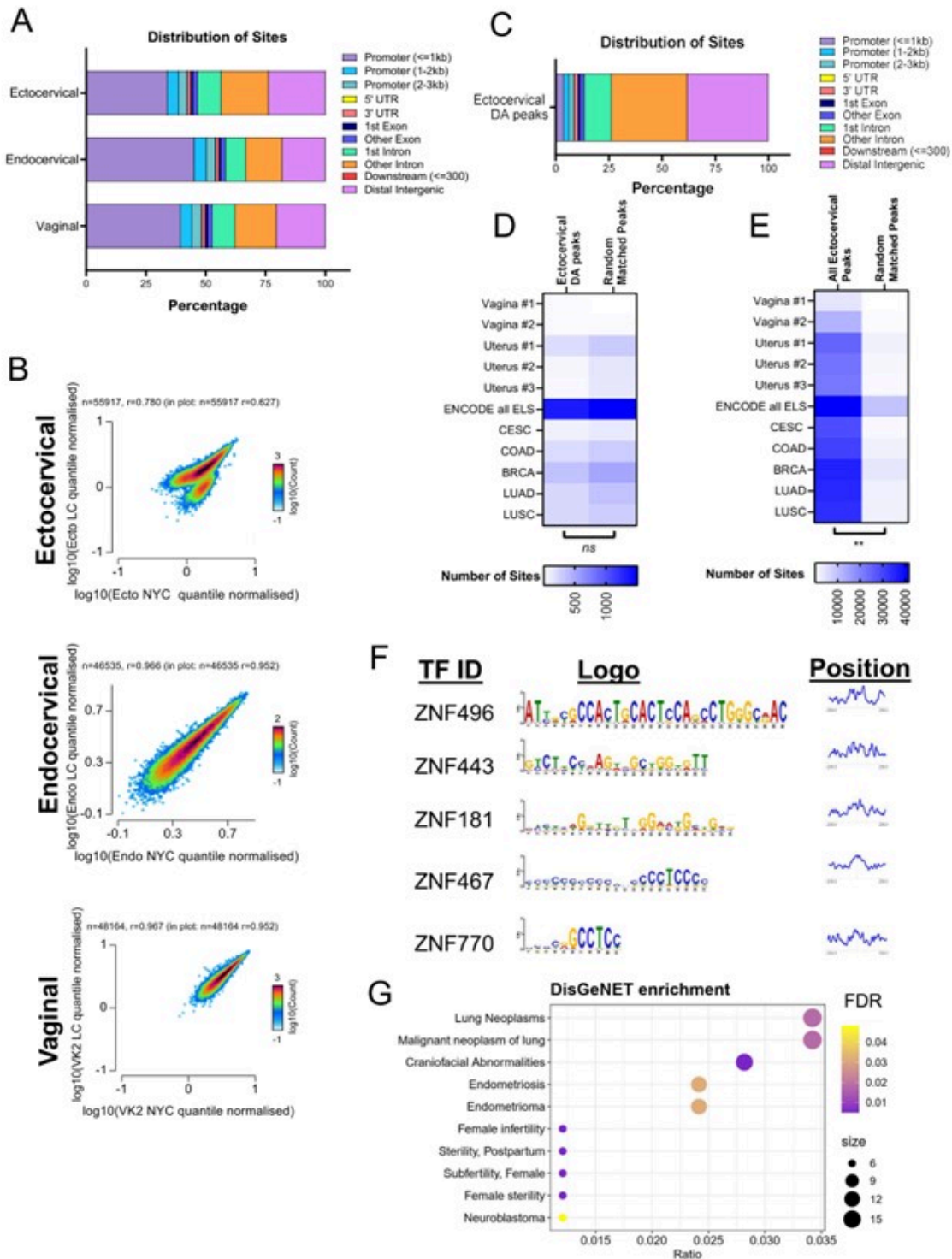
Figure 4

Chromatin accessibility was disrupted primarily in ectocervical cells exposed to *L. crispatus* supernatants. (A) Distribution of consensus sites by percentage. (B) Scatterplot of normalized counts between NYC and *L. crispatus* supernatant treatment by cell type-specific consensus peaks. (C) Distribution of differentially accessible sites in ectocervical cells across the genome. (D) Overlap of differentially accessible sites and matched number of random sites (E) all ectocervical peaks and

matched number of random sites with published sites from ENCODE and multiple primary cancer specimens. (F) Motif analysis of downregulated differentially accessible sites with motif logo and graph of positional distribution based on the center of the peak of the top 5 motifs. (G) Bubble chart of DisGeNET enrichment of transcription factors identified by the motif analysis. DA: Differential Accessibility; ELS: Enhancer-like signatures; CESC: Cervical Squamous Cell Carcinoma; COAD: Colon Adenocarcinoma; BRCA: Breast Invasive Carcinoma; LUAD: Lung Adenocarcinoma; LUSC: Lung Squamous Cell Carcinoma.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementalmethods.docx
- SupplementalTable1.xlsx
- SupplementalTable2.xlsx
- SupplementalTables35.docx
- SupplementalTable6af.xlsx
- SupplementalTable7.xlsx
- SupplementalTable8ac.xlsx
- SupplementalTable9.xlsx
- SupplementalFigures18.docx