International Journal of Women's Health

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Investigation of the Endometrial Microbiome in Recurrent Pregnancy Loss Individuals: Microbial Imbalance and Network Fragility

Bolun Zhang[1,2], Shaochong Lin[1,2], Sidong Wang[2,3], Weiyu Chen[4], Yushu Chen[4], Dandan Cao[2], Qingzhi Liu[2], Yuanqing Yao[1]

[1]Department of Medical College, Nankai University, Tianjin, People's Republic of China; [2]Center for Reproductive Medicine, Shenzhen Hospital, The University of Hong Kong, Shenzhen, People's Republic of China; [3]Department of Obstetrics and Gynaecology, Li ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, People's Republic of China; [4]Hangzhou Veritas Genetics Medical Institute Co., Ltd., Hangzhou, People's Republic of China

Correspondence: Yuanqing Yao, Department of Medical College, Nankai University, Tianjin, People's Republic of China, Email yaoyq_hkuszh@126.com

Purpose: Emerging evidence suggests that an abnormal endometrial microbiota may be a potential factor contributing to recurrent pregnancy loss (RPL). This study aimed to characterize the endometrial microbiota in patients with RPL and to explore its association with miscarriage.

Patients and Methods: Based on specific inclusion and exclusion criteria, EndoMetrial Microbiome Assay (EMMA) data from women attending clinics were collected and categorized into RPL and control groups according to their miscarriage history. Species diversity analysis, differential microbiota analysis, and machine learning methods were employed to identify key microbial genera associated with RPL. Microbial network analysis was then performed to further characterize the endometrial microbiome in patients with RPL.

Results: No significant differences in α-diversity were observed between the RPL and control groups across multiple indices (all $P > 0.05$); however, β-diversity differed significantly (Euclidean distance, $P = 0.039$). Regarding species composition, the control group showed a significantly higher abundance of *Lactobacillus*, whereas the RPL group had increased levels of pathogenic bacteria, including *Gardnerella, Staphylococcus*, and *Streptococcus*. Machine learning identified three key genera associated with RPL: *Streptococcus, Chryseobacterium*, and *Fusobacterium*. Microbial network analysis further revealed the fragility of the endometrial microbial community in patients with RPL.

Conclusion: These findings offer novel insights into the mechanisms of endometrial microenvironmental changes in patients with RPL and highlight potential microbial biomarkers and therapeutic targets for future clinical applications.

Keywords: recurrent pregnancy loss, endometrial microbiota, machine learning, microbial biomarkers

Introduction

Recurrent pregnancy loss (RPL) remains one of the most challenging conditions in reproductive medicine. It affects up to 5% of couples attempting conception[1] and is defined by the European Society of Human Reproduction and Embryology (ESHRE) as the occurrence of at least two pregnancy losses prior to 24 weeks of gestation.[2] Although several risk factors have been identified, including uterine abnormalities, endocrine dysfunction, autoimmune conditions, thrombotic disorders, and chromosomal anomalies, the etiology of RPL remains largely heterogeneous and incompletely understood. Among the proposed factors, infections and imbalances in the reproductive tract microbiota have been implicated in certain cases.[3] Pathogenic microorganisms and their endotoxins may elicit abnormal inflammatory responses in the endometrium, impair decidualization, and disrupt immune regulation at the maternal–fetal interface, potentially leading to abnormal embryonic development or early pregnancy loss.[4] However, no conclusive evidence supports a direct causal role of a specific microbial species in RPL,[5] likely due to complex interspecies interactions and host-specific factors that

Received: 12 April 2025
Accepted: 27 August 2025
Published: 3 September 2025

complicate such investigations. High-throughput sequencing technologies, particularly 16S rRNA sequencing, now allow more detailed characterization of microbial communities in human tissues and offer new perspectives on these interactions.[6]

Traditionally, the uterus was considered a sterile environment. This notion was challenged by Mitchell et al, who used qPCR and 16S rRNA sequencing to identify a complex microbial community within the uterine cavity, including *Lactobacillus* sp., *Acinetobacter* sp., and *Pseudomonas* sp.[7] This discovery has prompted increasing interest in the role of the endometrial microbiota in fertility.[8] Studies have shown that women with a *Lactobacillus* sp.-dominant microbiota (>90%) have significantly higher implantation, pregnancy, and live birth rates during IVF-ET than those with non-dominant microbiota.[9] Recent reviews emphasize a growing body of evidence linking endometrial microbial dysbiosis to adverse reproductive outcomes, particularly RPL.[10] Additionally, one study reported that elevated *Ureaplasma urealyticum* levels in the endometrium were associated with increased risks of miscarriage and preterm birth.[11]

Despite these advances, the comprehensive characteristics of the endometrial microbiota in RPL, including diversity patterns and community interactions, remain poorly understood. The microbiota is not a random collection of bacteria but plays a critical role in regulating immune homeostasis and maintaining the stability of the endometrial microenvironment through intricate ecological networks. Therefore, this study aimed to characterize the endometrial microbiome in patients with RPL, with specific objectives to investigate microbial alterations, identify potential microbial biomarkers, and construct ecological networks of microbial interactions.

Materials and Methods

Study Population

This retrospective cohort study was conducted using clinical and microbiome data from patients who underwent the EndoMetrial Microbiome Assay (EMMA) at the Reproductive Medicine Clinic of the University of Hong Kong–Shenzhen Hospital between June 1, 2023, and March 31, 2024. Patients were categorized into two groups based on clinical history: the RPL group and the control group. RPL was defined as two or more pregnancy losses occurring before 24 weeks of gestation, regardless of whether conception was spontaneous or assisted. Patients were excluded if they met any of the following criteria: (1) female age $\geq$ 40 years; (2) body mass index (BMI) outside the range of 18.5–30 kg/m²; (3) chromosomal abnormalities in either partner; and (4) incomplete clinical or microbiome data. Patients with pregnancy loss due to fetal chromosomal abnormalities confirmed by chorionic villus sampling were also excluded from the RPL group.

Of the 178 patients initially screened, 122 met the eligibility criteria and were included in the final analysis: 89 in the RPL group and 33 in the control group. A detailed flowchart illustrating participant selection is provided in Figure S1. The study protocol received approval from the Ethics Committee of the University of Hong Kong–Shenzhen Hospital (Ethical Approval No. [2024]375).

Clinical Data Collection

Clinical data were extracted from the hospital's electronic medical records. Variables included age (years), BMI (kg/m²), anti-Müllerian hormone (AMH, ng/mL), gravidity, parity, and the number of prior spontaneous abortions. To protect patient confidentiality, all data were anonymized and coded before analysis, in accordance with ethical guidelines.

Endometrial Sample Collection

Endometrial samples were collected during the mid-secretory phase of the menstrual cycle to ensure consistency in hormonal milieu and endometrial receptivity. A standardized protocol was followed to minimize contamination and variability. The vaginal area was first disinfected with povidone-iodine and then thoroughly rinsed with sterile saline to remove any residual antiseptics. Cervical discharge was absorbed using a sterile dry cotton swab. A sterile, single-use endometrial sampling device (Biopsy-Mistogy Tube; Shanghai Jiabao Medical Health Technology Co., Ltd., China) was then gently inserted into the uterine fundus. After withdrawing the inner core to generate negative pressure, the catheter

was rotated and maneuvered to aspirate endometrial tissue. Care was taken to avoid contact with the vaginal wall during insertion and withdrawal. All procedures were conducted under strict aseptic conditions.

EMMA

The EMMA was performed to profile the endometrial microbiota and assess whether the uterine environment was conducive to implantation and pregnancy. EMMA utilizes next-generation sequencing (NGS) of the 16S rRNA gene to identify and quantify bacterial DNA with high precision. Genomic DNA was extracted from endometrial tissue using the QIAsymphony DSP DNA Mini Kit (Cat. No. 937236), and DNA concentrations were measured via spectrophotometry (MultiSkan GO, Thermo Fisher). Sequencing libraries were prepared using the Ion Plus Fragment Library Kit (Thermo Fisher), and sequencing was conducted on the Ion GeneStudio S5 Prime platform (Thermo Fisher). Raw reads were quality-controlled and processed using the QIIME2 pipeline. Low-quality sequences were filtered, trimmed, and aligned to the SILVA 16S rRNA reference database for taxonomic annotation.

Bioinformatics and Statistical Analysis

Downstream analysis of 16S rRNA sequencing data was performed using the MicrobiotaProcess R package. This included assessment of α-diversity and β-diversity, sample distance calculations, comparisons of microbial composition across taxonomic levels, and identification of differentially abundant taxa.[12] To identify key genera significantly associated with RPL, three machine learning algorithms were applied: support vector machine (SVM) using the e1071 package, least absolute shrinkage and selection operator (LASSO) regression via the glmnet package, and random forest (RF) classification via the randomForest package. Genera consistently identified across models were used to construct a binary logistic regression model, and a nomogram was generated to visualize predictive outcomes. Model performance was assessed using receiver operating characteristic (ROC) curve analysis. Additionally, microbial co-occurrence networks were constructed and compared between groups using the ggClusterNet package.[13] Baseline characteristics were compared using the Mann–Whitney *U*-test. All analyses were conducted using R software (version 4.4.1).

Results

Baseline Characteristics of Patients

A total of 122 patients (89 in the RPL group and 33 in the control group) were included in this study. No statistically significant differences were found in age, BMI, AMH levels, or parity. However, compared with the control group, the RPL group had significantly higher gravidity and number of abortions (Table 1).

Comparison of Microbial Diversity

Before comparing microbial diversity, rarefaction curves were plotted for all samples (Figure S2a). The curves showed that the number of species approached saturation as the number of reads increased, indicating sufficient sequencing depth.

Table 1 Patient Baseline Characteristics

Group	Total	Control	RPL	P value
Enrolled Population	122	33	89	
Age (y)	32.99±3.62	33.03±3.60	32.98±3.65	0.89
BMI (kg/m²)	21.64±2.21	21.79±2.50	21.59±2.10	0.78
AMH (ng/mL)	3.16±2.36	2.24±2.13	2.76±2.45	0.381
Gravidity	2.53±1.47	0.85±0.57	3.15±1.18	<0.001
Parity	0.29±0.49	0.18±0.46	0.32±0.49	0.09
Abortions	2.22±1.31	0.64±0.49	2.81±0.10	<0.001

Abbreviations: RPL, Recurrent pregnancy loss; BMI, Body Mass Index; AMH, Anti-Mullerian Hormone.

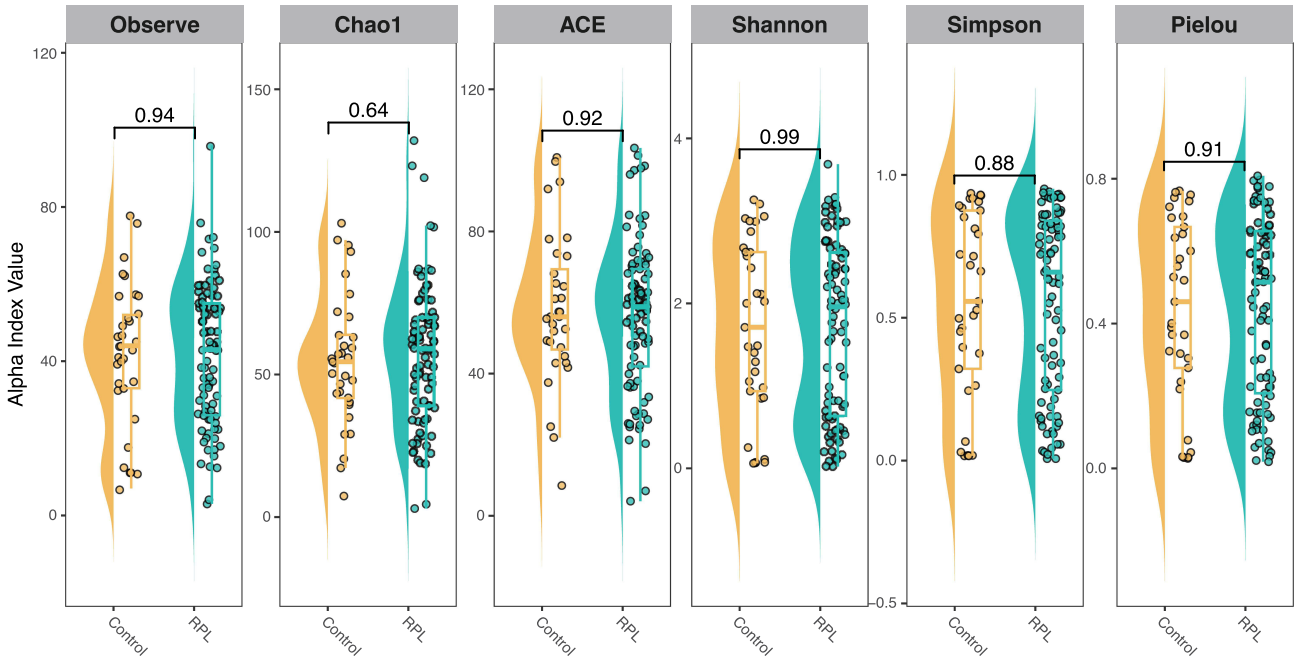

Figure 1 Alpha Diversity Indices in the Control and RPL Groups. Split violin plots overlaid with boxplots showing the distribution and variation of six alpha diversity indices (Chao1, ACE, Shannon, Simpson, and Pielou indices) in the control and RPL groups. Statistical comparisons between groups were performed using the Wilcoxon rank-sum test. *P*-values are shown for each pair. RPL: recurrent pregnancy loss.

Alpha Diversity

Alpha diversity measures species richness and evenness within individual samples. The rarefaction curves for both groups were highly similar (Figure S2b). Furthermore, no statistically significant differences were observed between the RPL and control groups in several alpha diversity indices, including Chao1, ACE, Shannon, Simpson, and Pielou indices (Figure 1).

Beta Diversity

Beta diversity assesses differences in species composition between samples or communities. Among the taxa identified, 48% were shared between the two groups, 12% were unique to the control group, and 40% were unique to the RPL group, indicating greater taxonomic uniqueness in the RPL group (Figure 2a).

Euclidean distances were calculated to measure variation in microbial composition in multidimensional space (Figure S3). In the control group, there were no statistically significant differences in Euclidean distances (P = 0.27). In contrast, the RPL group showed borderline significance (P = 0.052), indicating some heterogeneity. The overall variance in Euclidean distances between groups was statistically significant (P = 0.039), suggesting distinct microbial community composition (Figure 2b).

Principal coordinate analysis (PCoA) revealed that PCo1 and PCo2 accounted for 27.07% and 11.88% of the variance, respectively. Although some overlap was noted in the PCoA plot, a trend toward distinct separation was evident along the PCo1 axis, indicating differences in microbial community composition between the groups (Figure 2c).

Comparison of Microbial Abundance at Different Taxonomic Levels

At the phylum level, *Firmicutes* and *Proteobacteria* were the dominant phyla. The proportion of *Actinobacteria* was significantly increased in the RPL group, whereas the abundance of Bacteroidetes was notably reduced (Figure 3a).

At the class level, *Bacilli, Actinobacteria*, and *Betaproteobacteria* accounted for more than 80% of the microbial population. Among these, *Actinobacteria* was significantly elevated in the RPL group (Figure 3b).

Lactobacillales and *Actinomycetales* were more abundant in the control group, whereas *Bifidobacteriales* was more abundant in the RPL group. In addition, the RPL group contained a greater variety of other taxa (Figure 3c).

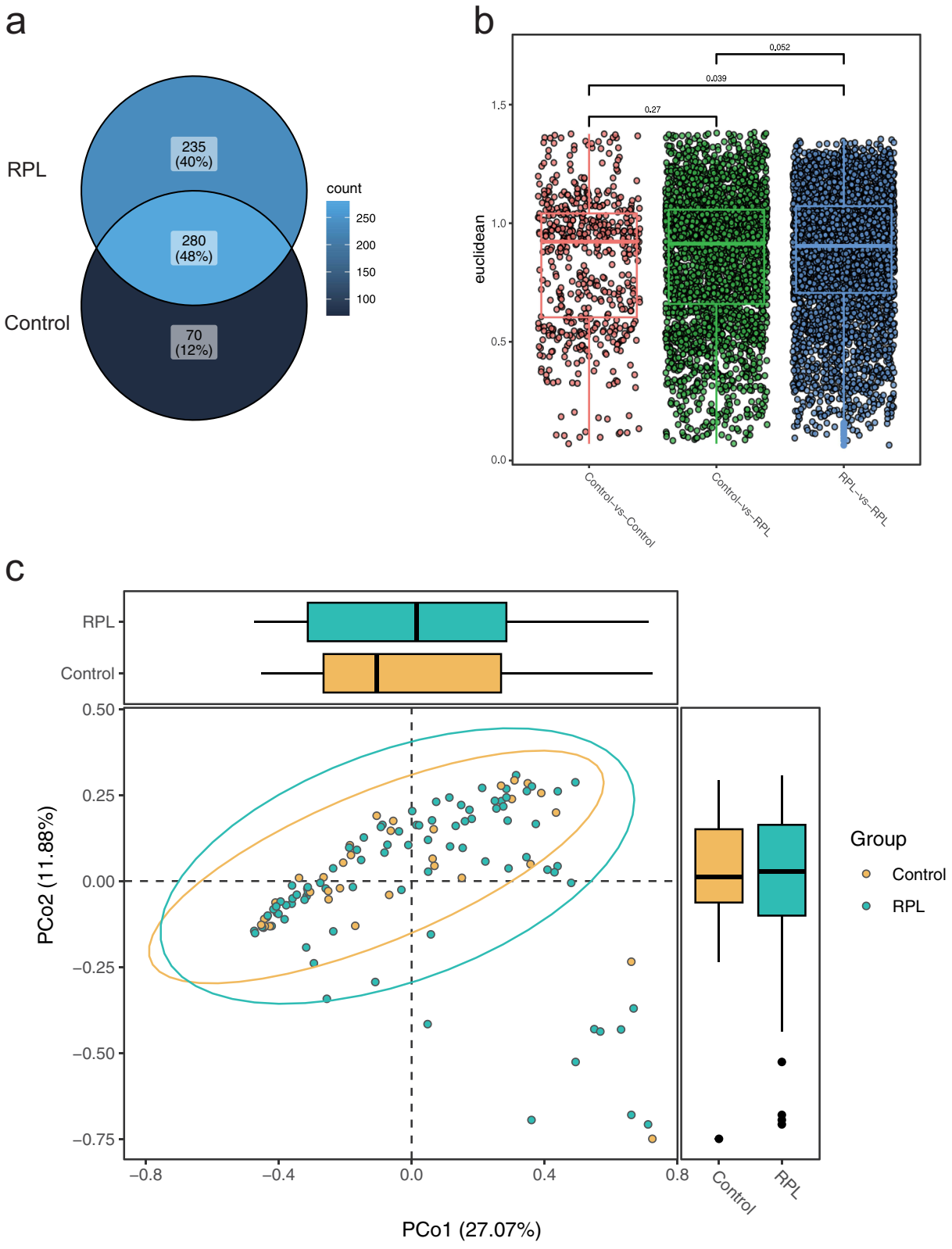

Figure 2 Beta Diversity and PCoA Analysis of the Control and RPL Groups. (**a**) Venn diagram showing the number of taxa identified in the two groups. (**b**) Beta diversity comparison based on Euclidean distance. (**c**) PCoA plot based on Euclidean distance.
Abbreviations: PCoA, principal coordinate analysis; RPL, recurrent pregnancy loss.

https://doi.org/10.2147/IJWH.S534065

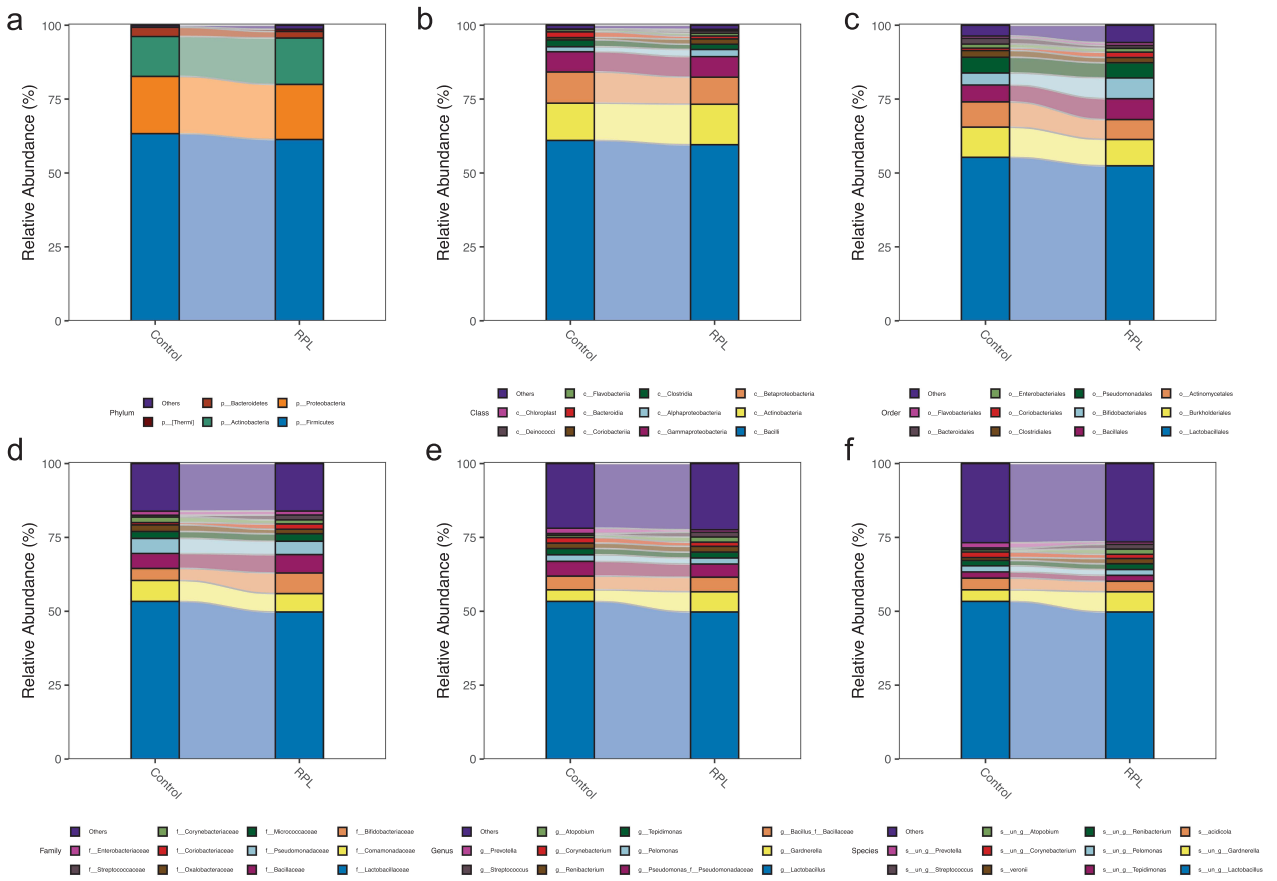

Figure 3 Taxonomic Composition at Different Levels Based on Relative Abundance. Stacked bar plots showing the relative abundance of microbial communities at six taxonomic levels—from phylum to species—in the control and RPL groups. Panels (**a**–**f**) represent the taxonomic levels: phylum, class, order, family, genus, and species, respectively.
Abbreviation: RPL, recurrent pregnancy loss.

At the family level, *Lactobacillaceae* was more abundant in the control group, whereas the abundance of *Bifidobacteriaceae* was higher in the RPL group (Figure 3d).

Due to limitations in taxonomic annotation at the species level, results at the genus and species levels were largely consistent. *Lactobacillus* was the dominant genus in the endometrial microbiome but was notably reduced in the RPL group. In contrast, *Gardnerella* and *Streptococcus* were significantly increased in the RPL group (Figure 3e and f).

Regarding absolute microbial abundance, all taxonomic levels showed higher absolute abundance in the RPL group than in the control group, suggesting an overall increase in microbial load in the endometrium of patients with RPL, possibly linked to the overgrowth of specific pathogens (Figure S4).

Differential Taxa Identification Across Multiple Taxonomic Levels

While stacked bar plots comparing relative abundances across taxonomic levels can miss low-abundance genera and lack statistical validation, we addressed this by identifying differential taxa across multiple levels, filtered for significance at $P < 0.05$, and quantified differences using linear discriminant analysis (LDA) scores (Figure 4). At the family level, *Cryomorphaceae, Streptococcaceae*, and *Xanthomonadaceae* were identified as differential taxa. Among them, *Streptococcaceae* and *Xanthomonadaceae* were significantly elevated in the RPL group, while *Cryomorphaceae* was enriched in the control group.

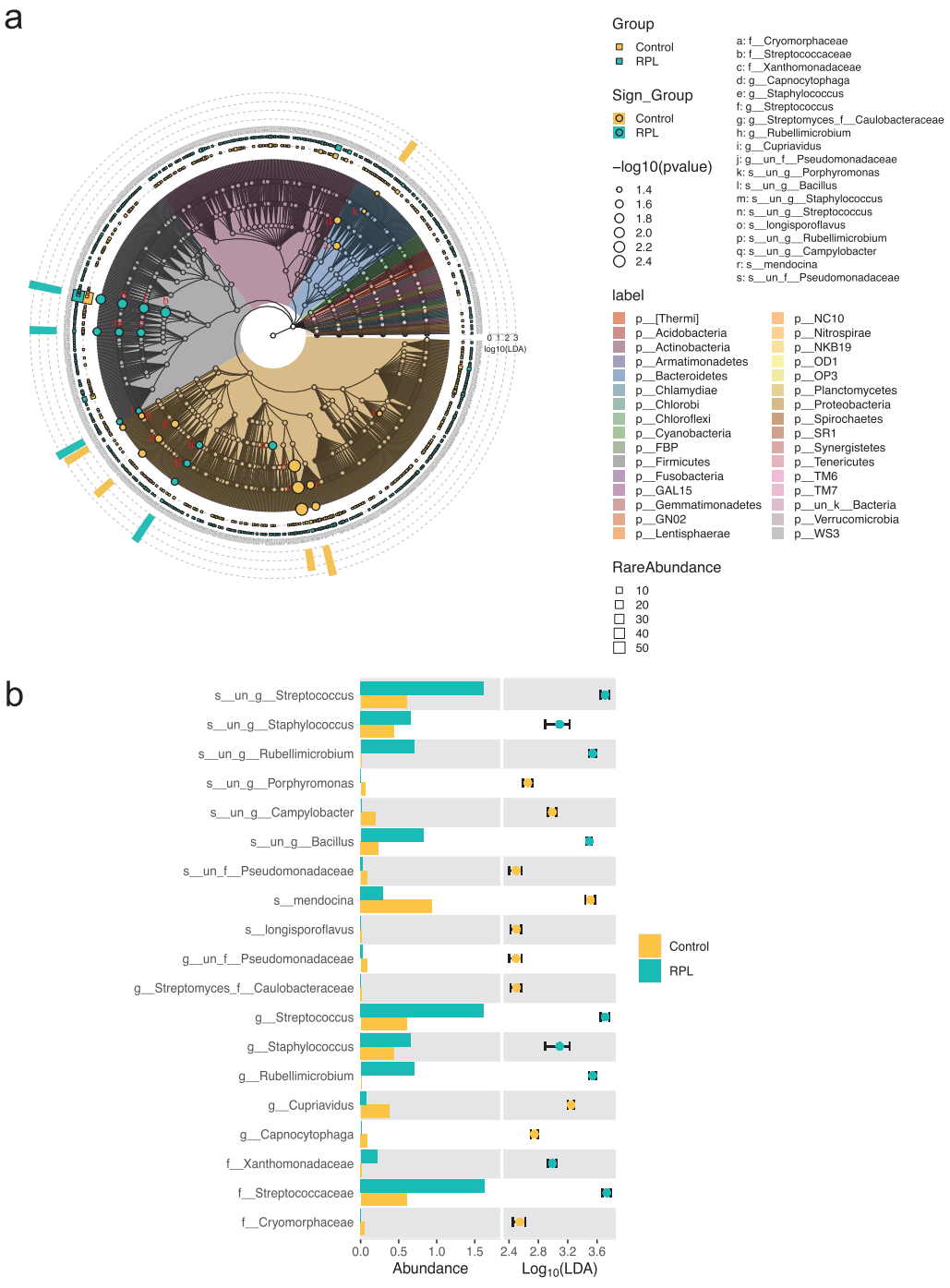

Figure 4 Differentially Abundant Taxa Between the Control and RPL Groups Across Multiple Taxonomic Levels. (**a**) Cladogram displaying the phylogenetic structure of the endometrial microbiota, organized hierarchically from the center (kingdom) to the outermost level (species), passing through phylum, class, order, family, and genus. The branches are colored based on taxonomic assignment at the phylum level to allow visual differentiation of microbial lineages. Differentially abundant taxa between the control and RPL groups were identified using LEfSe and are labeled (a–s) on the tree. These labels indicate taxa that were statistically significant ($P < 0.05$), and their positions in the tree show the taxonomic level at which differences emerge. The outermost layer of the cladogram displays bars representing the effect size (log10-transformed LDA scores) of each taxon. Green bars indicate taxa enriched in the RPL group, while yellow bars represent taxa enriched in the control group. (**b**) Comparative plots of differentially abundant taxa across multiple taxonomic levels between the control and RPL groups. The left panel shows the relative abundance of each taxon in both groups, while the right panel presents the corresponding effect size, measured as the log10-transformed LDA score. This score reflects the magnitude of difference between groups. Dots represent significantly different taxa ($P < 0.05$), with green indicating enrichment in the RPL group and yellow indicating enrichment in the Control group. Taxa not classified at the species level are labeled as "un" and are annotated using the nearest higher-level taxonomy.

Abbreviations: LDA, linear discriminant analysis; LEfSe, linear discriminant analysis effect size; RPL, recurrent pregnancy loss.

https://doi.org/10.2147/IJWH.S534065

In the phylogenetic tree shown in Figure 4a, 16 different taxa were identified at the genus and species levels, with several genera originating from the same family branches. Differential microbiota analysis revealed higher abundance of several potential pathogens in patients with RPL, including *Streptococcus, Staphylococcus, Pseudomonadaceae, Campylobacter*, and *Porphyromonas*. These genera are well-known for their association with various infectious diseases, and their enrichment may suggest a link to adverse reproductive outcomes.

Identification of RPL-Associated Genera Using Machine Learning Methods

Three machine learning methods were used for feature selection to identify RPL-associated genera. Figure 5a and b show the selection process in the LASSO model, where increasing the Lambda value reduced the coefficients of less important genera to zero, yielding a refined set of key genera. In the SVM model, the lowest error rate was observed with 20 selected features (Figure 5c). In the RF model, increasing the number of decision trees reduced the error rate (Figure 5d). The top five important genera identified were *Comamonas, Streptococcus, Bacillus, Lactobacillus*, and *Gardnerella* (Figure 5e).

The three models showed considerable overlap in the selected genera, confirming the robustness of the results. *Chryseobacterium, Fusobacterium*, and *Streptococcus* were consistently identified as key genera (Figure 5f). A logistic regression model based on these three genera achieved an AUC of 0.762, compared with AUC values of 0.645, 0.572, and 0.644 for *Streptococcus, Fusobacterium*, and *Chryseobacterium*, respectively, suggesting that the combined model significantly improved RPL prediction (Figure 5g).

The nomogram illustrated the contributions of *Streptococcus, Fusobacterium*, and *Chryseobacterium* to RPL. *Streptococcus* and *Chryseobacterium* were identified as risk factors, while *Fusobacterium* appeared protective. Genus values were mapped to scores, summed, and used to estimate individual RPL risk, thereby quantifying the link between genus abundance and RPL and offering a foundation for personalized treatment (Figure 5h).

Construction of the Endometrial Microbiota Network

The microbiome is not merely a sum of different microbial groups; microbial interactions are critical in shaping the endometrial microenvironment and local immune response. Interaction networks constructed using the top 300 genera by abundance revealed a more complex microbial interaction network in the control group than in the RPL group. In the control group, both cooperative and antagonistic interactions were observed, while only a few cooperative interactions were present in the RPL group (Figure 6a).

The microbial network in the control group exhibited distinct modular structures, with four module hubs and three connectors. In contrast, the RPL group lacked similar key nodes, indicating reduced network function (Figure 6b).

Network stability analysis showed that node connectivity was higher in the control group, and the network remained stable even after extensive node removal. Conversely, the RPL group exhibited low connectivity that declined sharply with node removal, suggesting that its microbial network was more fragile and susceptible to collapse (Figure 6c). As genera were gradually removed, the number of remaining taxa declined in both groups, but more sharply in the RPL group (Figure 6d). Removal of module hubs further accelerated species loss in the RPL network, underscoring its structural vulnerability (Figure 6e).

Within the top 50 genera, several highly connected modules were identified, with *Bacillus, Pontibacillus*, and *Pelomonas* as prominent central nodes. *Lactobacillus* also occupied a central role in the uterine microbial network, highlighting its importance in maintaining a healthy uterine environment. In the RPL group, the network was sparse, and *Lactobacillus* had no observable interactions with other genera, indicating a diminished protective role in patients with RPL (Figure S5).

Discussion

Although the role of microorganisms in human health and disease is receiving increasing attention, many unknowns remain regarding the microbiota under specific pathological conditions.[14] Recently, characteristics of the endometrial microbiome have been uncovered, particularly in the field of reproductive health.[15] Abnormal alterations in the endometrial microbiota are closely correlated with adverse reproductive outcomes. Certain potential pathogens, such as *Gardnerella* and *Atopobium*, have been linked to reproductive disorders such as chronic endometritis (CE) and implantation failure.[16–18] However, a comprehensive understanding of the endometrial microbiota in patients with RPL

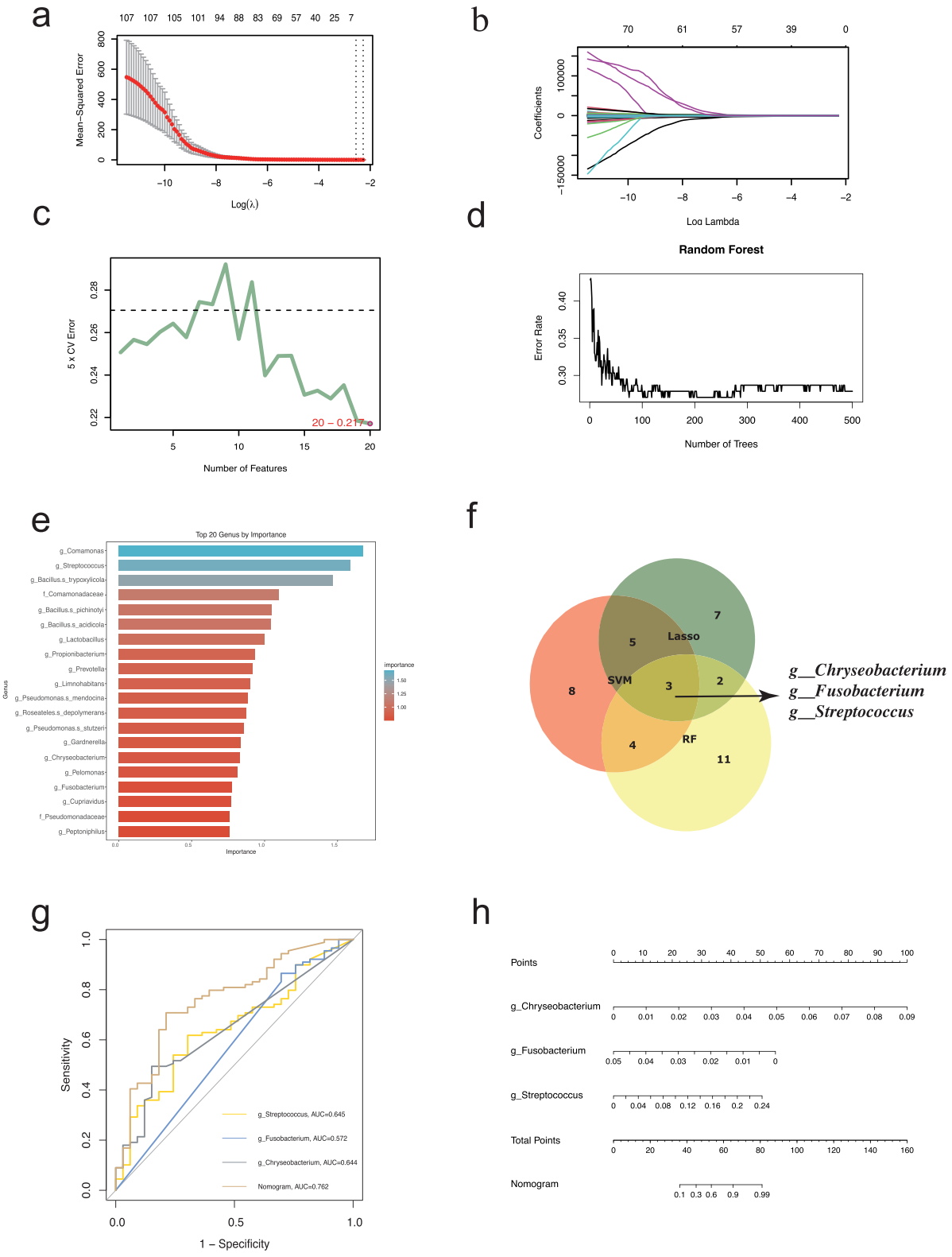

Figure 5 Identification of RPL-Associated Key Genera Through Multiple Machine Learning Methods. (**a** and **b**) Lasso regression was used to screen and identify significant genera. (**c**) SVM 5-fold cross-validation error rate plot. (**d**) RF cross-validation error rate plot. (**e**) Top 20 genera identified by RF, ranked by importance. (**f**) Venn diagram showing the overlap of important genera identified by three machine learning methods. (**g**) ROC curve of the regression model based on three key genera. (**h**) Nomogram of the model constructed from the three key genera.

Abbreviations: Lasso, least absolute shrinkage and selection operator; RF, random forest; ROC, receiver operating characteristic; RPL, recurrent pregnancy loss; SVM, support vector machine.

https://doi.org/10.2147/IJWH.S534065

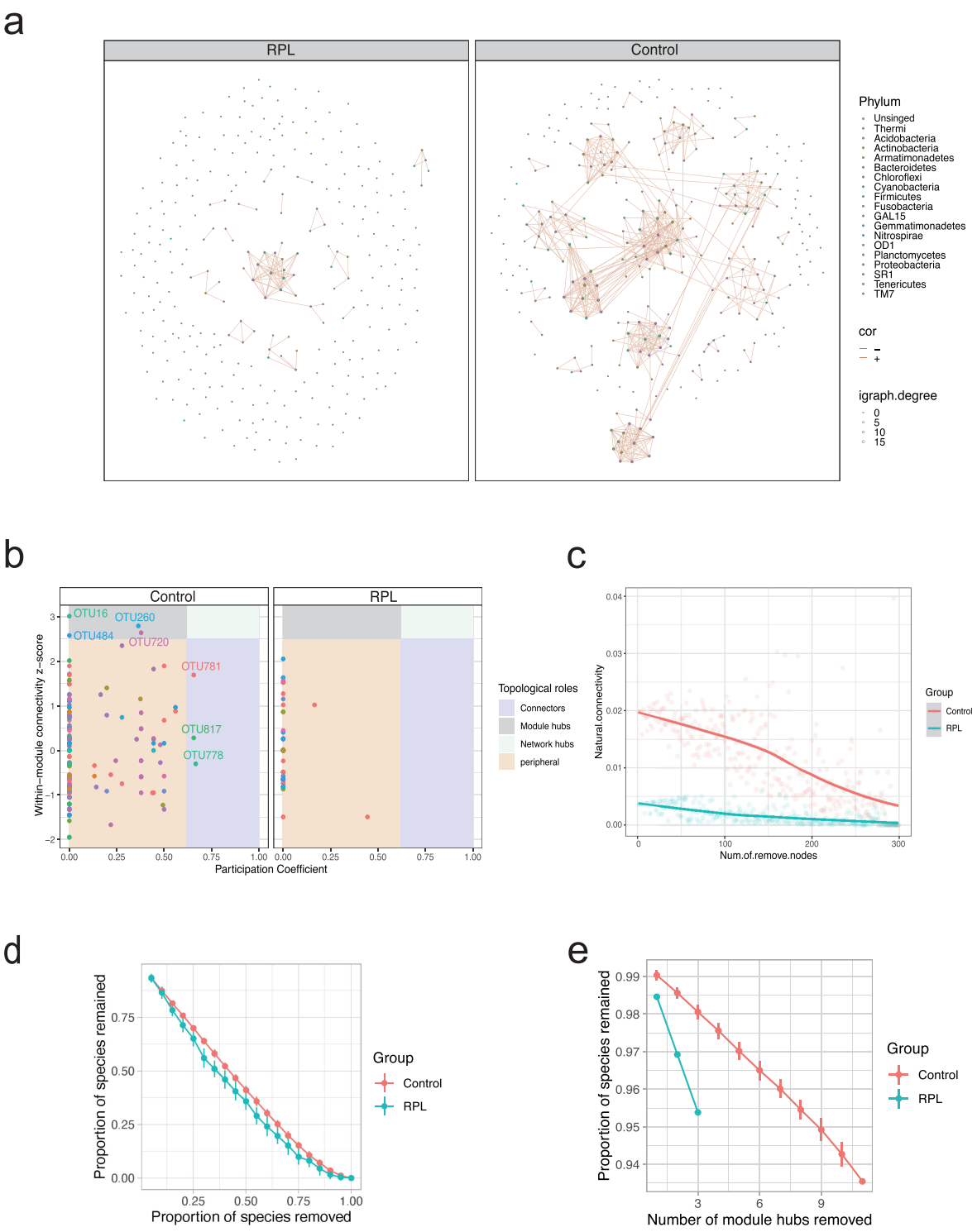

Figure 6 Construction of the Endometrial Microbiota Network. (**a**) Microbial interaction network constructed using the top 300 genera ranked by relative abundance. Each node represents a genus, with node colors indicating taxonomic affiliation at the phylum level. Edges represent microbial interactions: brown lines indicate positive (cooperative) associations, and blue lines indicate negative (competitive or antagonistic) associations. (**b**) Topological roles of nodes in the microbial co-occurrence network. Each dot represents an OTU positioned by its within-module connectivity (Z-score) and participation coefficient. Nodes are classified into four categories—peripherals, connectors, module hubs, and network hubs—based on their topological roles, with background shading indicating each role. The control group shows more module hubs and connectors than the RPL group, reflecting higher network complexity and connectivity. (**c**) Natural connectivity under progressive random node removal. The control network shows a slower decline in connectivity compared to the RPL network, indicating greater structural resilience. (**d**) Robustness analysis based on random node removal. Species loss accelerates more rapidly in the RPL network, reflecting lower ecological stability. (**e**) Robustness analysis after targeted removal of module hubs. Disruption of key taxa leads to a sharp collapse in the RPL network, whereas the control network remains more stable.

Abbreviation: RPL, recurrent pregnancy loss.

remains limited, especially regarding microbial diversity and functional interaction networks. Therefore, we conducted an in-depth analysis of EMMA data to systematically characterize the endometrial microbiota in patients with RPL, identify key genera, and construct microbial interaction networks to explore their potential association with RPL.

Comparative Analysis of Microbial Diversity Patterns

Species diversity is a key aspect of endometrial microbiome research. In the present study, multiple alpha diversity indices were not significantly different between the two groups, indicating consistent species richness and evenness within individual samples. This finding aligns with the findings of Lozano et al,[19] who also reported no significant differences in alpha diversity between the RIF and control groups. Similarly, another study found no difference in endometrial alpha diversity between patients with ongoing and non-ongoing pregnancies.[20] However, some studies have indicated that samples dominated by non-Lactobacillus species exhibit higher Shannon indices and that these patients tend to have poorer pregnancy outcomes.[9] These findings suggest that while alpha diversity is important for assessing the overall structure of microbial communities, its impact on pregnancy outcomes may depend on microbial composition and functional state.

Beta diversity measures differences in microbial communities between groups. In this study, we observed a statistically significant difference in beta diversity between the RPL and control groups, attributed to a higher number of unique taxa in the RPL group, consistent with previous findings.[19] Notably, within-group differences in the RPL group were also statistically significant, suggesting possible heterogeneity. Although there was some overlap in the PCoA analysis, the observed trend of separation indicated differences in microbial community composition between the groups, warranting further investigation.

Key Differential Microbial Taxa Between RPL and Control Groups

We next explored differences in microbial composition between groups at various taxonomic levels. Consistent with previous reports, *Lactobacillus* was the dominant genus in the endometrium.[21] In this study, the relative abundance of *Lactobacillales* (order level) was higher in the control group than in the RPL group, suggesting a protective role. Peuranpää et al reported a significant reduction in *Lactobacillus* abundance in patients with RPL,[22] and another study showed significantly higher implantation, pregnancy, and live birth rates during IVF-ET in women with *Lactobacillus*-dominated endometria.[9] These findings underscore the importance of *Lactobacillus* in maintaining reproductive health, and its depletion may lead to adverse pregnancy outcomes, particularly in RPL.

Homeostasis of the female reproductive tract microbiota relies on *Lactobacillus* dominance from the vagina and cervix to the uterine cavity. A decline in *Lactobacillus* abundance increases the risk of local infections and inflammatory responses.[23] First, lactic acid produced by *Lactobacillus* helps maintain an acidic environment in the reproductive tract, which inhibits pathogen invasion and colonization, thus protecting the reproductive system from potential harmful microorganisms.[24] Second, lactic acid possesses anti-inflammatory properties that help regulate the immune response in the reproductive tract. It can inhibit pro-inflammatory cytokine production via Toll-like receptors and promote IL-10 secretion, thereby reducing the cytotoxicity of natural killer (NK) cells.[25] Another study found that *Lactobacillus* and its metabolites regulate the immune status of the endometrium by upregulating the anti-inflammatory factor IL-1RA and suppressing pro-inflammatory signals.[26] Therefore, *Lactobacillus* not only effectively inhibits pathogens through its antibacterial and anti-inflammatory properties but also helps maintain a healthy microbial balance by modulating the local immune microenvironment.

In this study, a significant enrichment of *Actinobacteria* was observed in the RPL group, which may be associated with increased abundance of certain taxa, such as *Bifidobacteriaceae* and *Gardnerella*, at lower taxonomic levels in the RPL group. Cariati et al examined the distribution of endometrial microbiota in patients undergoing IVF-ET and reported that *Actinobacteria* appeared only in the non-pregnant group.[27] Another study revealed an increased abundance of *Actinobacteria* in the endometrium of non-pregnant patients,[28] suggesting a potential adverse effect of *Actinobacteria* on pregnancy. Reports of actinomycotic endometritis describe endometrial infiltration by neutrophils and plasma cells.[29] Additionally, two studies demonstrated increased *Actinobacteria* abundance in the endometrium of patients with inflammation-related diseases, including endometriosis and endometrial polyps, suggesting that it may act as an inducer

of immune responses, contributing to pathophysiological changes.[30,31] This immune activation could be the mechanism through which *Actinobacteria* affects pregnancy outcomes.

Gardnerella, an anaerobic bacterium linked to bacterial vaginosis, produces virulence factors that allow it to attach to host epithelial cells and form biofilms, which confer treatment resistance and protect other microorganisms, making infections more difficult to clear.[32] Here, we observed a significant increase in *Gardnerella* in the RPL group, consistent with previous findings.[22,33] It has been speculated that *Gardnerella* ascends from the vagina to the uterus, disrupting the immune balance of the endometrium and negatively affecting pregnancy.

Our differential microbiota analysis and machine learning results revealed a significant enrichment of *Streptococcus* in the endometrium of patients with RPL, suggesting its potential role in the occurrence of RPL. While some *Streptococcus* species colonize asymptomatically, enrichment has been linked to implantation failure.[34] Moreover, Group B *Streptococcus* (GBS) is associated with miscarriage and preterm premature rupture of membranes (PPROM) in late pregnancy[35] and may cause intrauterine infections and fetal death through vertical transmission.[36] Flaherty reported that GBS induces immune dysregulation and cell death by activating the mitogen-activated protein kinase (MAPK) and nuclear factor-κB (NF-κB) pathways, leading to miscarriage, preterm birth, and other pregnancy complications.[37] These findings reinforce the notion that *Streptococcus* acts as a key pathogen contributing to adverse pregnancy outcomes, implying that it may be an important factor in triggering changes in the endometrial microenvironment and immune dysregulation.

Pathogens such as *Escherichia coli, Streptococcus, Staphylococcus*, and *Mycoplasma* have been implicated in CE.[38] In line with this, we observed increased *Staphylococcus* and *Streptococcus* abundance in patients with RPL, suggesting that these individuals may be more susceptible to CE. Growing evidence has shown a strong association between CE and female infertility, RPL, and implantation failure. Imbalances in the endometrial microbiota and pathogenic bacterial overgrowth may lead to the abnormal infiltration of inflammatory cells, ultimately resulting in miscarriage or implantation failure.[39] In patients with unexplained RPL, imbalanced states in the endometrial microbiota, particularly the increase in pathogenic bacteria, may play a critical role in the mechanisms underlying their miscarriages.

In this study, we also identified some taxa that have been less frequently reported in prior studies, such as *Cupriavidus, Mendocina, Campylobacter*, and *Porphyromonas*, all of which were significantly more abundant in the control group, suggesting that these microorganisms may be essential for maintaining endometrial microbiota homeostasis. A study in mice detected *Campylobacter* in the uterus post-ovulation, suggesting that it may participate in reproductive regulation through metabolic pathways and steroid biosynthesis.[40] Another study found the presence of *Porphyromonas* in the endometrium but reported no association with inflammatory responses.[41] Jain et al revealed that *Porphyromonas* in the endometrium is associated with changes in *Lactobacillus* abundance and may be linked to endometrial cancer.[42]

We also observed increased *Bacillus* abundance in the RPL group. A study co-culturing primary endometrial epithelial cells with *Bacillus* showed that it could induce epithelial cell death within 24 h and significantly increase the expression of pro-inflammatory cytokines.[43] Another study also found that *Bacillus* can infect the placenta and fetal membranes, leading to adverse pregnancy outcomes such as miscarriage.[44] Chang et al demonstrated that intrauterine infusion of platelet-rich plasma (PRP) improved endometrial receptivity by altering the composition of endometrial microbiota, with the most notable microbial shift being a reduction in *Bacillus*.[45] These findings suggest that *Bacillus* may contribute to inflammation and pathogenic changes in the endometrium, adversely affecting pregnancy.

Feature Selection and Visualization of Core Microbiota via Machine Learning Approaches

This study is the first to use machine learning to identify core genera associated with RPL, namely *Chryseobacterium, Fusobacterium*, and *Streptococcus*. Incorporating their relative abundance allowed effective RPL classification. *Streptococcus*, which is enriched as a pathogen, may play a key role in RPL, whereas *Fusobacterium* showed an inverse trend, suggesting a protective role. Overall, the combination of these three genera adds value to the RPL. Research on the roles of *Chryseobacterium* and *Fusobacterium* in reproductive outcomes is limited; however, some studies have revealed

their potential association with endometrial-related diseases. One study suggested that *Fusobacterium* infection enhances the proliferation, adhesion, and migration of endometrial fibroblasts, contributing to endometriosis.[46] Another study found abnormal enrichment of *Chryseobacterium* and *Fusobacterium* in the microbiota of patients with CE and endometrial polyps. This suggests that these genera may influence the local immune environment of the endometrium, potentially affecting the reproductive health. Although the current data are limited, it is advisable to better understand how these genera will specifically contribute to conditions such as RPL in the future.[47]

In this study, *Fusobacterium* showed a reverse scoring trend, suggesting a different mechanism. Microbial pathogenicity is modulated by environmental factors, community dynamics, and host interactions. Under certain conditions, *Fusobacterium* may compete with or interact with other microbes, inhibit pathogenic processes, and exert protective effects.

Network Instability of Endometrial Microbiota in RPL

Microbial interactions are essential for regulating the endometrial immune microenvironment. Network analysis offers a powerful tool to elucidate these relationships and understand how microbial communities contribute to endometrial health through coordinated ecological effects.[48] In this study, the microbial network in the RPL group was sparse and fragile, lacking key module hubs and connectors, and exhibiting low natural connectivity. It rapidly collapsed when a few nodes were removed. This indicates that the microbial community in the RPL group was ecologically less stable than that in the control group, suggesting a link between RPL and the degradation of the endometrial microbial network structure. Low connectivity and fragility of the microbial community may exacerbate immune dysregulation in the endometrium, further impairing its ability to support embryo implantation and pregnancy maintenance. Therefore, changes in the network structure of the endometrial microbiota are not only markers of RPL but may also play a key role in its pathological process.[49]

Strengths and Limitations of the Study

This study has several strengths. It provides a detailed characterization of the endometrial microbiome in patients with RPL, encompassing species diversity, microbial abundance, and differential microbiota analysis. The innovative application of machine learning facilitated the identification of key genera strongly associated with RPL. Furthermore, microbial network analysis revealed structural fragility within the endometrial microbiota of patients with RPL, offering new insights into potential microbe-related mechanisms. However, this study also has limitations. As a retrospective analysis, it lacks prospective cohort recruitment, which may introduce selection bias and limit the generalizability of the findings. To mitigate this, strict and consistent inclusion and exclusion criteria were applied. In addition, due to the retrospective design, no environmental or vaginal control samples were collected. Although all samples were obtained using standardized aseptic techniques to minimize contamination, the absence of negative controls remains a limitation. The sample size also limits the ability to explore potential associations between endometrial microbiota features and the occurrence of RPL, underscoring the need for further investigation in large-scale, multicenter studies. Lastly, external factors such as environment, diet, and lifestyle may influence the endometrial microbiome. Future studies should incorporate these variables and employ integrated multi-omics approaches to better elucidate the microbial features and potential mechanisms linked to RPL.

Conclusion

This study presents an in-depth profiling of the endometrial microbiome in women with RPL, revealing significant differences compared to healthy controls. Protective bacteria, such as *Lactobacillus* spp., were reduced in patients with RPL, whereas pathogenic bacteria, including *Gardnerella* spp., *Streptococcus* spp., and *Staphylococcus* spp., exhibited increased diversity and abundance. Using machine learning, we identified core genera associated with RPL, particularly *Streptococcus* spp., *Chryseobacterium* spp., and *Fusobacterium* spp., which may contribute to miscarriage. Additionally, the endometrial microbial network in patients with RPL appeared more fragile and prone to disruption, suggesting that miscarriage in these patients may be largely correlated with microbial imbalance.

Data Sharing Statement

All raw data are available from the corresponding author or the first author upon reasonable request.

Ethics Approval and Informed Consent

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the University of Hong Kong–Shenzhen Hospital (protocol code Ethics [2024]375, approved on November 29, 2024). Written informed consent was obtained from all participants prior to EMMA testing, including consent for the use of sequencing data and clinical information for scientific research and publication.

Consent for Publication

All patients who underwent EMMA in this study signed informed consent forms and agreed to have their sequencing results and related information available for scientific research.

Acknowledgments

We would like to thank the assistant team at the Reproductive Medicine Center of the University of Hong Kong-Shenzhen Hospital, including Meilan Cao and Jiaman Li, for their support throughout this study. We also extend our gratitude to Hangzhou Yizhen Medical Laboratory for providing the EMMA technology, commercially known as HerAnova Endometrial Microbiome Analysis (HerBiota-Complete).

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This research was funded by the National Key Research and Development Program of China (Grant Nos. 2022YFC2702500, 2022YFC2702503), the Shenzhen Fundamental Research Program of China (Grant No. JCYJ20220818103013028), Guangdong Basic and Applied Basic Research Foundation, China (Grant No. 2023A1515220177), High-Level Hospital Program, Health Commission of Guangdong Province, China (Grant No. HKUSZH202207001). This work was also supported by the Shenzhen Basic Research General Project (grant no. JCYJ20230807113014029).

Disclosure

The author(s) report no conflicts of interest in this work.

References

1. Stephenson M, Kutteh W. Evaluation and management of recurrent early pregnancy loss. *Clin Obstet Gynecol*. 2007;50(1):132–145. doi:10.1097/GRF.0b013e31802f1c28
2. Bender Atik R, Christiansen OB, Elson J, et al. ESHRE guideline: recurrent pregnancy loss: an update in 2022. *Hum Reprod Open*. 2023;2023(1):hoad002. doi:10.1093/hropen/hoad002
3. Akbari S, Shahsavar F, Karami R, Yari F, Anbari K, Ahmadi SAY. Recurrent spontaneous abortion (RSA) and maternal KIR genes: a comprehensive meta-analysis. *JBRA Assist Reprod*. 2020;24(2):197–213. doi:10.5935/1518-0557.20190067
4. Liu Y, Wong KK, Ko EY, et al. Systematic comparison of bacterial colonization of endometrial tissue and fluid samples in recurrent miscarriage patients: implications for future endometrial microbiome studies. *Clin Chem*. 2018;64(12):1743–1752. doi:10.1373/clinchem.2018.289306
5. Galang RR, Chang K, Strid P, et al. Severe coronavirus infections in pregnancy: a systematic review. *Obstet Gynecol*. 2020;136(2):262–272. doi:10.1097/AOG.0000000000004011
6. Dominguez-Bello MG, Godoy-Vitorino F, Knight R, Blaser MJ. Role of the microbiome in human development. *Gut*. 2019;68(6):1108–1114. doi:10.1136/gutjnl-2018-317503
7. Mitchell CM, Haick A, Nkwopara E, et al. Colonization of the upper genital tract by vaginal bacterial species in nonpregnant women. *Am J Obstet Gynecol*. 2015;212(5):611.e611–619. doi:10.1016/j.ajog.2014.11.043

8. Franasiak JM, Scott RT. Reproductive tract microbiome in assisted reproductive technologies. *Fertil Steril*. 2015;104(6):1364–1371. doi:10.1016/j.fertnstert.2015.10.012
9. Moreno I, Codoñer FM, Vilella F, et al. Evidence that the endometrial microbiota has an effect on implantation success or failure. *Am J Obstet Gynecol*. 2016;215(6):684–703. doi:10.1016/j.ajog.2016.09.075
10. Odendaal J, Black N, Bennett PR, Brosens J, Quenby S, MacIntyre DA. The endometrial microbiota and early pregnancy loss. *Hum Reprod*. 2024;39(4):638–646. doi:10.1093/humrep/dead274
11. Shi Y, Yamada H, Sasagawa Y, Tanimura K, Deguchi M. Uterine endometrium microbiota and pregnancy outcome in women with recurrent pregnancy loss. *J Reprod Immunol*. 2022;152:103653. doi:10.1016/j.jri.2022.103653
12. Xu S, Zhan L, Tang W, et al. MicrobiotaProcess: a comprehensive R package for deep mining microbiome. *Innovation*. 2023;4(2):100388. doi:10.1016/j.xinn.2023.100388
13. Wen T, Xie P, Yang S, et al. ggClusterNet: an R package for microbiome network analysis and modularity-based multiple network layouts. *Imeta*. 2022;1(3):e32. doi:10.1002/imt2.32
14. Peterson J, Garges S, Giovanni M, et al. The NIH human microbiome project. *Genome Res*. 2009;19(12):2317–2323. doi:10.1101/gr.096651.109
15. Davies J. In a map for human life, count the microbes, too. *Science*. 2001;291(5512):2316. doi:10.1126/science.291.5512.2316b
16. Johnston-MacAnanny EB, Hartnett J, Engmann LL, Nulsen JC, Sanders MM, Benadiva CA. Chronic endometritis is a frequent finding in women with recurrent implantation failure after in vitro fertilization. *Fertil Steril*. 2010;93(2):437–441. doi:10.1016/j.fertnstert.2008.12.131
17. Yang R, Du X, Wang Y, Song X, Yang Y, Qiao J. The hysteroscopy and histological diagnosis and treatment value of chronic endometritis in recurrent implantation failure patients. *Arch Gynecol Obstet*. 2014;289(6):1363–1369. doi:10.1007/s00404-013-3131-2
18. Cicinelli E, Matteo M, Tinelli R, et al. Prevalence of chronic endometritis in repeated unexplained implantation failure and the IVF success rate after antibiotic therapy. *Hum Reprod*. 2015;30(2):323–330. doi:10.1093/humrep/deu292
19. Lozano FM, Lledó B, Morales R, et al. Characterization of the endometrial microbiome in patients with recurrent implantation failure. *Microorganisms*. 2023;11(3):741. doi:10.3390/microorganisms11030741
20. Franasiak JM, Werner MD, Juneau CR, et al. Endometrial microbiome at the time of embryo transfer: next-generation sequencing of the 16S ribosomal subunit. *J Assist Reprod Genet*. 2016;33(1):129–136. doi:10.1007/s10815-015-0614-z
21. Punzón-Jiménez P, Labarta E. The impact of the female genital tract microbiome in women health and reproduction: a review. *J Assist Reprod Genet*. 2021;38(10):2519–2541. doi:10.1007/s10815-021-02247-5
22. Peuranpää P, Holster T, Saqib S, et al. Female reproductive tract microbiota and recurrent pregnancy loss: a nested case-control study. *Reprod Biomed Online*. 2022;45(5):1021–1031. doi:10.1016/j.rbmo.2022.06.008
23. Abou Chacra L, Fenollar F, Diop K. Bacterial vaginosis: what do we currently know? *Front Cell Infect Microbiol*. 2021;11:672429. doi:10.3389/fcimb.2021.672429
24. Tachedjian G, Aldunate M, Bradshaw CS, Cone RA. The role of lactic acid production by probiotic Lactobacillus species in vaginal health. *Res Microbiol*. 2017;168(9–10):782–792. doi:10.1016/j.resmic.2017.04.001
25. Ilhan ZE, Łaniewski P, Thomas N, Roe DJ, Chase DM, Herbst-Kralovetz MM. Deciphering the complex interplay between microbiota, HPV, inflammation and cancer through cervicovaginal metabolic profiling. *EBioMedicine*. 2019;44:675–690. doi:10.1016/j.ebiom.2019.04.028
26. Delgado-Diaz DJ, Tyssen D, Hayward JA, Gugasyan R, Hearps AC, Tachedjian G. Distinct immune responses elicited from cervicovaginal epithelial cells by lactic acid and short chain fatty acids associated with optimal and non-optimal vaginal microbiota. *Front Cell Infect Microbiol*. 2019;9:446. doi:10.3389/fcimb.2019.00446
27. Cariati F, Carotenuto C, Bagnulo F, et al. Endometrial microbiota profile in in-vitro fertilization (IVF) patients by culturomics-based analysis. *Front Endocrinol*. 2023;14:1204729. doi:10.3389/fendo.2023.1204729
28. Kong Y, Liu Z, Shang Q, et al. The disordered vaginal microbiota is a potential indicator for a higher failure of in vitro fertilization. *Front Med*. 2020;7:217. doi:10.3389/fmed.2020.00217
29. Chiesa-Vottero AG. Actinomycotic Endometritis. *Int J Gynecol Pathol*. 2019;38(2):138–142. doi:10.1097/PGP.0000000000000476
30. Wessels JM, Domínguez MA, Leyland NA, Agarwal SK, Foster WG. Endometrial microbiota is more diverse in people with endometriosis than symptomatic controls. *Sci Rep*. 2021;11(1):18877. doi:10.1038/s41598-021-98380-3
31. Medina-Bastidas D, Camacho-Arroyo I, García-Gómez E. Current findings in endometrial microbiome: impact on uterine diseases. *Reproduction*. 2022;163(5):R81–r96. doi:10.1530/REP-21-0120
32. Schwebke JR, Muzny CA, Josey WE. Role of Gardnerella vaginalis in the pathogenesis of bacterial vaginosis: a conceptual model. *J Infect Dis*. 2014;210(3):338–343. doi:10.1093/infdis/jiu089
33. Moreno I, Simon C. Deciphering the effect of reproductive tract microbiota on human reproduction. *Reprod Med Biol*. 2019;18(1):40–50. doi:10.1002/rmb2.12249
34. Moreno I, Garcia-Grau I, Perez-Villaroya D, et al. Endometrial microbiota composition is associated with reproductive outcome in infertile patients. *Microbiome*. 2022;10(1):1. doi:10.1186/s40168-021-01184-w
35. Rocchetti TT, Marconi C, Rall VL, Borges VT, Corrente JE, da Silva MG. Group B streptococci colonization in pregnant women: risk factors and evaluation of the vaginal flora. *Arch Gynecol Obstet*. 2011;283(4):717–721. doi:10.1007/s00404-010-1439-8
36. American College of Obstetricians and Gynecologists. Prevention of Group B Streptococcal Early-Onset Disease in Newborns: ACOG Committee Opinion, Number 797. *Obstet Gynecol*. 2020;135(2):e51–e72. doi:10.1097/AOG.0000000000003668
37. Flaherty RA, Magel M, Aronoff DM, Gaddy JA, Petroff MG, Manning SD. Modulation of death and inflammatory signaling in decidual stromal cells following exposure to group B streptococcus. *Infect Immun*. 2019;87(12). doi:10.1128/IAI.00729-19
38. Kitaya K, Yasuo T. Commonalities and disparities between endometriosis and chronic endometritis: therapeutic potential of novel antibiotic treatment strategy against ectopic endometrium. *Int J Mol Sci*. 2023;24(3):2059. doi:10.3390/ijms24032059
39. Chen P, Chen P, Guo Y, Fang C, Li T. Interaction between chronic endometritis caused endometrial microbiota disorder and endometrial immune environment change in recurrent implantation failure. *Front Immunol*. 2021;12:748447. doi:10.3389/fimmu.2021.748447
40. Pu S, Wang M, Wang J, et al. Metagenomic analysis reveals a dynamic microbiome with diversified adaptive functions that respond to ovulation regulation in the mouse endometrium. *BMC Genomics*. 2023;24(1):615. doi:10.1186/s12864-023-09712-8
41. Karstrup CC, Klitgaard K, Jensen TK, Agerholm JS, Pedersen HG. Presence of bacteria in the endometrium and placentomes of pregnant cows. *Theriogenology*. 2017;99:41–47. doi:10.1016/j.theriogenology.2017.05.013

42. Jain M, Mladova E, Dobychina A, et al. Comparison of microbial profiles and viral status along the vagina-cervix-endometrium continuum of infertile patients. *Syst Biol Reprod Med*. 2023;69(4):310–319. doi:10.1080/19396368.2023.2195040
43. Gärtner MA, Peter S, Jung M, Drillich M, Einspanier R, Gabler C. Increased mRNA expression of selected pro-inflammatory factors in inflamed bovine endometrium in vivo as well as in endometrial epithelial cells exposed to Bacillus pumilus in vitro. *Reprod Fertil Dev*. 2016;28(7):982–994. doi:10.1071/RD14219
44. Agerholm JS, Jensen NE, Dantzer V, Jensen HE, Aarestrup FM. Experimental infection of pregnant cows with Bacillus licheniformis bacteria. *Vet Pathol*. 1999;36(3):191–201. doi:10.1354/vp.36-3-191
45. Chang Y, Peng J, Zhu Y, et al. How platelet-rich plasma (PRP) intra-uterine injection improve endometrial receptivity of intrauterine adhesions in women: a time-series-based self-controlled study. *J Reprod Immunol*. 2023;156:103796. doi:10.1016/j.jri.2023.103796
46. Muraoka A, Suzuki M, Hamaguchi T, et al. Fusobacterium infection facilitates the development of endometriosis through the phenotypic transition of endometrial fibroblasts. *Sci Transl Med*. 2023;15(700):eadd1531. doi:10.1126/scitranslmed.add1531
47. Liang J, Li M, Zhang L, et al. Analysis of the microbiota composition in the genital tract of infertile patients with chronic endometritis or endometrial polyps. *Front Cell Infect Microbiol*. 2023;13:1125640. doi:10.3389/fcimb.2023.1125640
48. Baker JM, Chase DM, Herbst-Kralovetz MM. Uterine microbiota: residents, tourists, or invaders? *Front Immunol*. 2018;9:208. doi:10.3389/fimmu.2018.00208
49. Molina NM, Sola-Leyva A, Saez-Lara MJ, et al. New opportunities for endometrial health by modifying uterine microbial composition: present or future? *Biomolecules*. 2020;10(4):593. doi:10.3390/biom10040593