



PhD program in Biomedicine and Biotechnology

**“Deciphering the functional impact of the endometrial
microbiome in reproductive failure”**

Author:

Bruno Toson

Supervised by:

Dr. Carlos Simón Vallés

Dr. Inmaculada Moreno Gimeno

Valencia, July 2024



D. Carlos Antonio Simón Vallés, Catedrático del Dpto. de Pediatría, Obstetricia y Ginecología de la Universitat de València

y

Dña. Inmaculada Moreno Gimeno, Investigadora Principal en la Fundación Carlos Simón para la Investigación en Salud de la Mujer e Instituto de Investigación Sanitaria INCLIVA

CERTIFICAN:

Que la presente memoria, titulada "*Deciphering the functional impact of the endometrial microbiome in reproductive failure*", corresponde al trabajo realizado bajo su dirección por D. Bruno Toson, para su presentación como Tesis Doctoral en el Programa de Doctorado en Biomedicina y Biotecnología de la Universitat de València.

Y para que conste firman el presente certificado en Valencia, a 18 de julio de 2024.

SIMON
VALLES
CARLOS
ANTONIO -
73757585N

Firmado
digitalmente por
SIMON VALLES
CARLOS ANTONIO
- 73757585N
Fecha: 2024.07.19
10:00:41 +02'00'

Fdo. Carlos Antonio Simón Vallés

MORENO
GIMENO
INMACULADA -
48387153Y

Firmado digitalmente
por MORENO GIMENO
INMACULADA -
48387153Y
Fecha: 2024.07.18
15:32:31 +02'00'

Fdo. Inmaculada Moreno Gimeno

PUBLICATIONS AND CONGRESS ASSISTANCE

The work presented in this doctoral thesis was carried out at Carlos Simon Foundation/INCLIVA Biomedical Research Institute thanks to the financial support of the grant “Santiago Grisolia GRISOLIAP/2021/176” from the Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital.

Throughout the doctoral thesis, work supported by this grant generated results presented in the following scientific publications and congresses:

Original Articles:

1. Bellver, J., Gonzalez-Monfort, M., González, S., **Toson, B.**, Labarta E., Castellón, G., Mariani, G., Vidal, C., Giles, J., Cruz, F., Ballesteros, A., Ferrando M., García-Velasco, J. A., Valbuena, D., Vilella, F., Parras, M., Tercero-Atencia, E., Simon, C., Moreno I. Reproductive BioMedicine Online; under review.
2. **Toson, B.**, Gonzalez-Monfort, M., Perez-Villaroya, D., Yue, Y., Yin, J., Vilella, F., Moreno, I., Simón, C. Altered taxonomic and functional bacterial networks in the endometrium associate with differential transcriptomics in women with reproductive failure. Under preparation.

Reviews and Book chapters:

1. **Toson, B.**, Simon, C., & Moreno, I. (2022). The Endometrial Microbiome and Its Impact on Human Conception. International Journal of Molecular Sciences, 23(1), 485. <https://doi.org/10.3390/ijms23010485>

2. Moreno, I., Garcia-Grau, I., González, M., **Toson, B.** Microbiota and Pathogen Screening in the Female Reproductive Tract. Encyclopedia of Reproduction, Third Edition. Accepted, in print.

Scientific meetings:

1. **Work title:** Altered taxonomic and functional bacterial networks in the endometrium of women with reproductive failure.

Authors: **Toson, B.**, Gonzalez-Monfort, M., Perez-Villaroya, D., Yue, Y., Yin, J., Vilella, F., Moreno, I., Simón, C.

Conference: Innovations in Reproductive Medicine 2030: Technologies in IVF (RM2030) congress. November 16-18, 2023, in Paris, France

Type of presentation: Oral

2. **Work title:** Deciphering the endometrial host-microbiome interactions associated with reproductive health.

Authors: Gonzalez-Monfort, M., **Toson, B.**, Perez-Villaroya, D., Yue, Y., Yin, J., Vilella, F., Moreno, I., Simón, C.

Conference: Innovations in Reproductive Medicine 2030: Technologies in IVF (RM2030) congress. November 16-18, 2023, in Paris, France

Type of presentation: Oral

3. **Work title:** Development and Characterization of a 3D Human Endometrium-on-a-Chip Model

Authors: Zaragozano, S., Moncayo-Arlandi, J., **Toson, T.**, González-Fernández, J., Pablo Grases, J., Pardo-Figuerez, M., Ochando, A., Simon, C., Vilella, F.

Conference: 71st SRI Annual Scientific Meeting. March 12-16, 2024, in Vancouver, Canada

Type of presentation: Oral

4. **Work title:** Development and Characterization of a 3D Human Endometrium-on-a-Chip Model

Authors: Zaragozano, S., Moncayo-Arlandi, J., **Toson, T.**, González-Fernández, J., Pablo Grases, J., Pardo-Figuerez, M., Ochando, A., Simon, C., Vilella, F.

Conference: 3D Organ Modelling in Health and Disease. April 8-9, 2024, in Leuven, Belgium

Type of presentation: Poster

AGRADECIMENTOS/ AGRADECIMIENTOS

Agradeço à toda minha família pelo suporte ao longo de toda a minha vida. Em especial ao meu núcleo familiar (pais, irmã e avó), mas principalmente à minha mãe, Carla, que é a grande responsável por cada uma das minhas conquistas. Sem ela, nada disso teria sido possível.

Aos meus melhores amigos do Brasil, Amanda, Ana, Duda, Isadora, Izabela, Joana, Luísa, Marlana e Uriel, por tornarem a vida tão mais bonita. Também deixo aqui um agradecimento especial àqueles que tiveram uma participação essencial durante o meu doutorado e foram minha maior fonte de força e tranquilidade nessa jornada difícil: Andreo, Cristina, Iuri, José, Maria, Matheus, Nathan e Rafinha. Minha sorte por ter vocês é imensa, e palavras nunca poderiam descrever tudo o que vocês significam para mim.

Aos meus (muitos!) antigos colegas de trabalho, por acreditarem em mim e tornarem minha caminhada tão engrandecedora. Aos meus anteriores mentores e chefes Maria, Rafael Elias, Carol, Rafael Roesler e Aïda. Em especial ao Tiago e ao Zeca, que me deram a primeira oportunidade de iniciação científica e trouxeram à vida o pesquisador que sempre existiu em mim.

A mis amigos de España, por todo el soporte y por enseñarme tantas cosas de vuestra preciosa cultura. Un saludo especial a mis compis Jaime, Maca, Vicent y Víctor; y a mis chiquis Alex, Andrea, Vivi y Zé: sois la familia que he encontrado en el otro lado del mundo. Me habéis ayudado a buscar solución para todo, pero también

me habéis hecho disfrutar mucho de esta experiencia y habéis cambiado mi vida para siempre.

A mis compañeros de la Fundación, por todos los momentos compartidos a lo largo de estos tres años: Adrián, Alba, Albichuela, Ana M., Ángela(s), Anna, Carla, Carolina, Diego, Esther, Fany, Irene, Jaime, Javi, María, María Cristina, Mari Carmen, Marta Gálvez, Moncayo, Neus, Nerea, Raúl, Sara, Sheila, Tery, Valli y Yolanda. Un agradecimiento especial a mi rubita Sofía, por todas las horas compartidas en el coche, además de muchas risas y confidencias también. A las miocovs Esther, Paulita y Sandra, por los incontables memes y apoyo todos los días, incluso por adoptarme como miocov honorario. Y por último, pero no menos a Ana Ochando, a quien debo muchísimo más de lo que las palabras jamás podrían decir. Juntos hemos trabajado mucho, pero también hemos construido una conexión que llevaré para siempre.

A mis compañeros Marta González y David Pérez, por todo el soporte durante la tesis y por todo vuestro empeño en construir un trabajo tan importante. Sin vuestro gran conocimiento, esto no habría sido posible.

A mis tutores de estancia, Dr. Daniel Ramón y Alba Arevalo, por haberme recibido tan bien y por todo el conocimiento compartido en el período que estuve en I2SysBio.

A mis directores de tesis Dr. Carlos Simón y Dra. Inmaculada Moreno, por su orientación y supervisión durante este proceso tan importante. Gracias por haberme acogido y haberme enseñado tanto sobre el mundo de la investigación.

A todos vocês, minha imensa gratidão por fazerem de mim quem eu sou.

RESUMEN

1. Introducción

1.1 El microbioma humano

El término microbiota define al conjunto de microorganismos que coexisten en un entorno definido, mientras que microbioma considera todo el hábitat, incluyendo los microorganismos, sus genomas y las condiciones circundantes (Marchesi & Ravel, 2015). En los últimos años se ha identificado la presencia de una microbiota propia en diferentes tejidos y órganos del cuerpo humano, como el intestino, la cavidad oral, el tracto genital femenino/masculino y la piel (El-Sayed et al., 2021a; Hou et al., 2022).

Dada la importancia de la microbiota en la salud humana, alteraciones en su composición pueden conducir a estados conocidos como disbiosis que se han asociado con numerosas patologías, incluyendo asma, dermatitis atópica, periodontitis, obesidad, y cáncer (El-Sayed et al., 2021a; Manos, 2022). No obstante, a pesar del creciente conocimiento sobre la influencia de la composición de la microbiota en la salud humana, poco se sabe sobre el impacto funcional de las comunidades y genomas microbianos en el origen de enfermedades. Por lo tanto, los nuevos avances en microbiología molecular y el uso de estudios multiómicos se centran ahora en los mecanismos por los cuales los microorganismos pueden desencadenar o facilitar enfermedades humanas.

1.1.1 Herramientas para detección del microbioma

El uso de la secuenciación del gen bacteriano que codifica para el ARNr 16S ha implicado una mejora significativa en los estudios de la microbiota al permitir la caracterización de comunidades completas, en contraposición al cultivo donde se aíslan los distintos componentes de una comunidad. Además, nuevas técnicas como la secuenciación completa de metagenomas y la secuenciación de ARN han facilitado el estudio de los genes presentes y funciones activas del microbioma, pero también han llevado a una mejor clasificación de las comunidades bacterianas a nivel de especie e incluso son capaces de llegar al nivel de cepa (Galloway-Peña & Hanson, 2020; Wensel et al., 2022).

La secuenciación completa de metagenomas permite la reconstrucción de la comunidad y el perfil funcional del microbioma a través de la secuenciación de fragmentos aleatorios de todo el ADN presente en una muestra. Mientras que la metagenómica proporciona una vista estática de la composición genética de las comunidades microbianas, la metatranscriptómica evalúa los patrones de expresión génica activos, permitiendo estudios sobre las actividades metabólicas en tiempo real entre el microbioma y el huésped. Al examinar qué genes se transcriben activamente y de manera dinámica, la metatranscriptómica revela información crucial sobre el repertorio funcional de las comunidades microbianas y cómo responden a los estímulos ambientales y las interacciones huésped-microorganismo (Ojala et al., 2023).

1.1.2 Microbioma del trato genital femenino y su impacto en la salud

Las bacterias del género *Lactobacillus* representan más del 90% del recuento bacteriano en el tracto genital de mujeres en edad reproductiva, y cuatro especies parecen ser dominantes: *Lactobacillus crispatus*, *L. iners*, *L. gasseri* y *L. jensenii* (Ravel et al., 2011). Además, se ha demostrado que el microbioma uterino es altamente diverso pero poco abundante en comparación con el del tracto genital inferior (Chen et al., 2017). Varios estudios han descrito el microbioma uterino en mujeres asintomáticas en edad reproductiva, entre el que destaca también una dominancia de especies del género *Lactobacillus* (Moreno et al., 2016, Kyono, Hashimoto, Nagai, et al., 2018).

La disbiosis del tracto reproductivo femenino se ha asociado con varios problemas ginecológicos y complicaciones obstétricas, como la vaginosis bacteriana, la endometritis crónica, la endometriosis, y el parto prematuro (Punzón-Jiménez & Labarta, 2021). Además, alteraciones en la composición de la microbiota endometrial también se asocian a infertilidad y peores resultados reproductivos en reproducción asistida, como fallo recurrente de implantación, aborto clínico y embarazo bioquímico (Moreno et al., 2022, Cariatì et al., 2023, Odendaal et al., 2024). Por lo tanto, la evaluación y modulación de la microbiota se ha propuesto como un enfoque factible para mejorar los resultados clínicos durante la terapia de reproducción asistida (ART).

1.1.3 Impacto funcional del microbioma en el huésped

Se ha demostrado de manera consistente que las bacterias pueden modular la homeostasis del tejido humano. Por ejemplo, productos derivados de la microbiota como transcritos, moléculas pequeñas, y proteínas pueden desencadenar la

expresión diferencial de genes y la producción diferencial de proteínas en el intestino del huésped (Nichols & Davenport, 2021). Además, se ha demostrado que las vesículas extracelulares de bacterias pueden transportar factores de virulencia y ARN de interferencia cortos a las células del huésped, interrumpiendo vías de señalización del hospedador y modulando su respuesta inmune (Charpentier et al., 2023). También se ha demostrado que ciertas bacterias pueden secretar metabolitos con la capacidad de unirse a receptores acoplados a proteínas G (GPCRs) del hospedador, culminando en una respuesta celular que puede ir desde cambios en la expresión génica hasta alteraciones en el metabolismo celular o el potencial de membrana (Oliveira de Souza et al., 2021). Sin embargo, estas interacciones no son necesariamente patogénicas, ya que las interacciones del hospedador con bacterias comensales son necesarias para el correcto funcionamiento del organismo y ayudan a mantener la homeostasis en distintas regiones del cuerpo.

Aunque muchas de esas interacciones se han demostrado en el intestino, se hipotetiza que mecanismos similares puedan ocurrir en el endometrio. El microbioma endometrial comensal puede participar en una relación simbiótica con el endometrio y sus mediadores inmunitarios locales, facilitando potencialmente la implantación y modulando la tolerancia inmune que es crucial para el éxito del embarazo. Sin embargo, cambios en la composición del microbioma endometrial podrían conducir a la desregulación de genes en las células endometriales necesarias para la consecución y mantenimiento del embarazo, afectando la homeostasis y contribuyendo a la infertilidad (Odendaal et al., 2024).

1.1.3.1 Resistencia a antibióticos

El resistoma comprende el grupo de genes que confieren resistencia a los antibióticos, e incluye genes que actúan a través de varios mecanismos para disminuir la eficacia de los medicamentos, como (i) la disminución de la permeabilidad celular y la capacidad de expulsar los antibióticos, (ii) las modificaciones de la molécula del antibiótico por acetilación, fosforilación o adenilación, (iii) la producción de β -lactamasas, o (iv) la modificación, amplificación o protección de los sitios diana de los antibióticos (Munita y Arias, 2016).

El uso de antibióticos juega un papel crucial en diferentes aspectos de la medicina reproductiva, sirviendo para prevenir y tratar infecciones que pueden afectar la fertilidad y la salud reproductiva (Cheng et al., 2022; Dalby y Stoner, 2022; Tomás et al., 2020). Hallazgos recientes demuestran que las intervenciones terapéuticas basadas en la detección de patógenos mediante secuenciación del gen que codifica para el ARNr 16S en muestras endometriales mejoran significativamente los resultados del embarazo en mujeres sometidas a tratamiento de infertilidad (Iwami et al., 2023). Por lo tanto, los antibióticos son herramientas valiosas en medicina reproductiva, y su uso consciente es esencial para prevenir el desarrollo de resistencia a los antibióticos que podría afectar significativamente los resultados clínicos de las pacientes. Sin embargo, es importante considerar que ciertos antibióticos pueden actuar indiscriminadamente sobre las bacterias beneficiosas, lo que potencialmente perturbaría el delicado equilibrio del microbioma. Por lo tanto, es esencial avanzar en el desarrollo de nuevas terapias antimicrobianas para abordar el aumento de resistencias, garantizando así la efectividad continua de los tratamientos y preservando la salud reproductiva de los pacientes.

1.1.3.2 Factores de virulencia

Los factores de virulencia son características específicas de cada microorganismo que mejoran su capacidad para establecer una colonización. Estos pueden actuar de distintas maneras, incluyendo la capacidad de introducirse en el huésped mediante exoenzimas, de evadir las defensas del huésped usando moduladores inmunes, de adquirir hierro y nutrientes del entorno a través de transportadores, de percibir cambios ambientales con el sistema de dos componentes y *quorum sensing*, de adherirse a superficies utilizando fimbrias o induciendo la formación de *biofilms*, entre otros. En conjunto, la presencia de factores de virulencia expresados por bacterias está estrechamente asociada con la capacidad del patógeno para establecer y mantener la infección en un huésped susceptible (Casadevall y Pirofski, 2009).

Teniendo en cuenta el problema actual de la resistencia/tolerancia a los antibióticos, el desarrollo de los antimicrobianos de nueva generación (NGAs) que tenga como diana estos mecanismos de virulencia representa una oportunidad para el tratamiento de enfermedades infecciosas. La estrategia de inhibir los factores de virulencia en lugar de la viabilidad microbiana tiene un inmenso potencial en la salud humana y animal, y es probable que el uso de NGAs en vacunas o terapias antimicrobianas tenga una importancia clínica cada vez mayor en el futuro. Además, el tratamiento con NGAs podría hacer que los microorganismos sean más vulnerables a ser eliminados por el sistema inmunológico y potencialmente volverlos más susceptibles a los antibióticos tradicionales (Gadar y McCarthy, 2023).

1.2 Respuesta del hospedador frente al microbioma

El hospedador también juega un papel crucial en la composición del microbioma, influyendo en la abundancia y diversidad de las comunidades microbianas en distintas zonas del cuerpo. Por ejemplo, se ha descrito que la composición de la microbiota intestinal puede ser inducida por el huésped a través de células dendríticas, o también secreción de anticuerpos IgA y péptidos antimicrobianos (AMPs) para ayudar a mantener la función barrera de la mucosa (Zheng et al., 2020). Además, los receptores de reconocimiento de patrones como los receptores tipo Toll (TLRs, de *Toll-like receptors*) ayudan en la defensa contra patógenos y regulan la abundancia de microorganismos comensales, manteniendo la integridad del tejido (Rakoff-Nahoum et al., 2004) y facilitando la respuesta inmune mediante la liberación de citocinas inflamatorias y anticuerpos (Burgueño y Abreu, 2020). Así, se hipotetiza que las variaciones en la regulación inmunológica dentro de la cavidad uterina, la cual contiene un ambiente inmunológico altamente activo y complejo, podrían tener un fuerte impacto en la microbiota endometrial (Gholiof et al., 2022).

1.3 Aplicaciones en medicina reproductiva

Debido a la creciente comprensión de la influencia de las bacterias en la fertilidad (Moreno et al., 2022), la modulación de la composición del microbioma endometrial por el uso de antibióticos o probióticos ha ganado fuerza en medicina reproductiva (Iwami et al., 2023, Kadogami et al., 2020). En concreto, entender el impacto de los factores bacterianos en el tracto genital femenino es esencial para comprender los mecanismos que contribuyen a la infertilidad y otras condiciones obstétricas, ya que pueden modular el metabolismo celular del huésped, el éxito de la colonización y la eficacia del tratamiento. En última instancia, el conocimiento de estas interacciones

bacteria-hospedador posibilitará el desarrollo de tratamientos innovadores que vayan más allá de los enfoques tradicionales de antibióticos y probióticos, ofreciendo estrategias terapéuticas novedosas para mejorar los resultados reproductivos.

2. Hipótesis

Dadas las evidencias sobre la relevancia de la microbiota endometrial en medicina reproductiva, se hipotetiza que las bacterias dentro de la cavidad uterina podrían ejercer una influencia significativa sobre las células del huésped y el entorno uterino en general. Por lo tanto, proponemos que el microbioma uterino, tanto en condiciones normales como disbióticas, contribuye a nivel molecular y celular a la fisiopatología uterina, creando un nicho que pueda facilitar o dificultar la implantación embrionaria y un embarazo saludable.

3. Objetivos

El objetivo general es caracterizar molecularmente la composición del microbioma de la cavidad uterina a nivel de especie y su impacto funcional en resultados reproductivos. Para esto, se establecen los siguientes objetivos específicos:

- I. Caracterizar los principales perfiles bacterianos de la microbiota endometrial **a nivel de especie** y su asociación con el éxito o fracaso reproductivo mediante **metagenómica** del líquido endometrial.
- II. Caracterizar las **funciones bacterianas** enriquecidas, además de factores de virulencia y genes asociados a resistencia a antibióticos, en distintos resultados reproductivos.

- III. Evaluar el impacto del microbioma uterino sobre el endometrio en pacientes con éxito o fracaso reproductivo mediante **metatranscriptómica** del endometrio.

4. Materiales y Métodos

4.1 Pacientes y muestras

Las muestras para este estudio se seleccionaron de una investigación previamente publicada por nuestro grupo que analizaba la asociación entre la composición de la microbiota endometrial y los resultados reproductivos en pacientes sometidas a técnicas de reproducción asistida (Moreno et al., 2022).

Después de recibir una transferencia personalizada de embriones en el momento de máxima receptividad endometrial siguiendo el análisis de receptividad endometrial (ERA), 27 mujeres lograron un resultado exitoso (nacimiento vivo, LB, 34,2%), mientras que 52 presentaron fracaso reproductivo: 38 no quedaron embarazadas (NP, 48,1%), 4 tuvieron un embarazo bioquímico (BP, 5%), 9 tuvieron un aborto clínico (CM, 11,4%), y 1 desarrolló un embarazo ectópico (EP, 1,3%). Todas las participantes proporcionaron consentimiento informado por escrito y fueron clasificadas en los siguientes grupos de acuerdo con sus resultados reproductivos (Zegers-Hochschild et al., 2017):

I. **Éxito reproductivo**, se refiere específicamente a la ocurrencia de un **nacimiento vivo**.

II. **Fracaso reproductivo**, comprende un espectro de resultados que no resultan en la consecución de un recién nacido vivo incluyendo:

- Embarazo bioquímico (BP) es una forma temprana de fracaso reproductivo, caracterizada por la detección temprana de la gonadotropina coriónica humana beta (β -hCG) en suero u orina, sin desarrollo embrionario posterior.
- Aborto clínico (CM) implica la pérdida de un embarazo intrauterino antes de las 22 semanas completas de edad gestacional.
- Embarazo ectópico (EP) ocurre cuando el óvulo fertilizado se implanta fuera del útero.
- Finalmente, la ausencia de embarazo (NP) incluye a aquellas pacientes que no tuvieron resultados positivos de β -hCG, ni saco gestacional visible después del examen ecográfico.

4.2 Metagenómica funcional

La preparación de las librerías de ADN para evaluación de la composición taxonómica y funcional en muestras de fluido endometrial se hizo siguiendo instrucciones de la casa comercial (Illumina) con un *target* de 150M de lecturas por muestra. Las lecturas pasaron por filtros de calidad y de longitud utilizando PRINSEQ (Schmieder & Edwards, 2011) y se fusionaron utilizando la herramienta de software FLASH (Magoč & Salzberg, 2011). Se eliminaron las lecturas del huésped utilizando el alineador Burrows-Wheeler Aligner (BWA) mapeado contra el genoma humano de referencia (Li & Durbin, 2010). Se eliminaron del análisis los géneros bacterianos que no colonizan humanos o que están consistentemente asociados con contaminantes de reactivos (Barton et al., 2006; Glassing et al., 2016; Grahn et al., 2003; Hashimoto & Kyono, 2019; Kitaya et al., 2019; Kyono, Hashimoto, Nagai, et al., 2018; Lauder et al., 2016; Laurence et al., 2014; Stinson et al., 2019; Tanner et al., 1998; Weyrich et

al., 2019). Además, aunque algunos estudios han detectado *Streptomyces* spp. en muestras clínicas (Gras et al., 2022), las especies pertenecientes a este género se encuentran principalmente en el suelo, el agua y las plantas, y por lo tanto también se excluyeron del análisis.

Los perfiles taxonómicos y funcionales se obtuvieron utilizando el análisis de alineación de Diamond (Buchfink et al., 2021) utilizando una base de datos de secuencias de referencia no redundante del NCBI (Pruitt, 2005). Las abundancias taxonómicas y funcionales se estimaron mediante el análisis de los resultados de salida de BLAST. Finalmente, los resultados taxonómicos y funcionales se procesaron utilizando el software estadístico R (R Core Team, 2018) para la descripción estadística y la representación gráfica del perfil taxonómico y funcional de las muestras.

Los tipos de estado de comunidad (CST) se obtuvieron utilizando el clustering jerárquico sobre la distancia euclidiana entre la composición de taxones de abundancia relativa de muestras. Se utilizó el software estadístico R (R Core Team, 2018) para la representación estadística y gráfica.

El análisis de redes de co-ocurrencia se realizó aplicando el algoritmo de correlación SparCC (Friedman & Alm, 2012) para fusionar recuentos brutos de anotación taxonómica y funcional de resultados metagenómicos. La implementación del algoritmo se realizó utilizando la biblioteca SpiecEasi R (Kurtz et al., 2015). Las redes taxonómicas se visualizaron con la biblioteca igraph (Csardi & Nepusz, 2006) en el entorno de programación R para análisis estadístico (R Core Team, 2018), considerando taxones con coeficientes de correlación superiores a 0.7. Las

asociaciones taxonómicas-funcionales se visualizaron con las bibliotecas *igraph* y *corrplot* (Wei et al., 2021).

Para el análisis de factores de virulencia se utilizó la Base de Datos de Factores de Virulencia (VFDB) (Chen, 2005) junto con el alineador de secuencias *Diamond* para búsquedas de proteínas y ARNs traducidos (Buchfink et al., 2021). Los perfiles de resistencia a antibióticos se obtuvieron mediante el mapeo con la Base de Datos Integral de Resistencia a Antibióticos (CARD) y la herramienta bioinformática de alineación del Identificador de Genes de Resistencia (RGI) (Alcock et al., 2023). En ambos casos, las abundancias diferenciales entre los resultados reproductivos se informaron utilizando la biblioteca *DESeq2* (Love et al., 2014) en el entorno de programación estadística R (R Core Team, 2018).

4.3 Metatranscriptómica

La preparación de las librerías de ARN para evaluación del transcriptoma en muestras de biopsia endometrial se hizo siguiendo instrucciones de la casa comercial (Illumina) con un *target* de 100M de lecturas por muestra. Las lecturas generadas tras la secuenciación fueron filtradas por calidad y anotadas utilizando el paquete de análisis de metatranscriptomas *SAMSA2* (Westreich et al., 2018). Para el alineamiento, primero se mapearon todas las lecturas contra el genoma humano utilizando *BWA* (Li & Durbin, 2009). Posteriormente, se empleó *SAMtools* para procesar los archivos de alineación (.bam) y se descartaron las lecturas no mapeadas (Li et al., 2009). Las fracciones ribosomales de las lecturas bacterianas fueron eliminadas manualmente. Finalmente, los resultados taxonómicos y funcionales de *SAMSA2* fueron procesados utilizando el software estadístico R (R Core Team, 2018).

para la descripción estadística y la representación gráfica del perfil taxonómico y funcional de la muestra.

4.4 Análisis estadístico

Se realizaron pruebas exactas de Fisher y pruebas de proporciones individuales utilizando el software estadístico R (R Core Team, 2018) para evaluar la correlación entre los CST y los resultados reproductivos. Los conjuntos de datos de secuenciación de metagenomas se analizaron en busca de taxones y funciones diferencialmente abundantes utilizando DESeq2 (Love et al., 2014) y LEfSe (Segata et al., 2011). Para el análisis de la transcriptómica humana, se utilizó la biblioteca edgeR de R para la detección de expresión génica diferencial (Robinson et al., 2010), y se utilizó la biblioteca clusterProfiler de R para el enriquecimiento de conjuntos de genes con términos de ontología génica (Wu et al., 2021).

En todos los casos, los valores significativos se consideraron cuando el $\log_2\text{FoldChange}$ (FC) era ≥ 2.0 o ≤ -2.0 y sus valores de p-valor ajustados (p-adj) o tasa de descubrimiento falso (FDR) eran < 0.05 .

5. Resultados y Discusión

5.1 Análisis taxonómico: microbiota

Utilizando análisis metagenómicos para llegar al nivel de especie, observamos CSTs similares a los encontrados en muestras vaginales (Gajer et al., 2012; Ravel et al., 2011), pero con algunas diferencias. En nuestro estudio, CST-I y -III mostraron una composición similar a la observada en vagina con dominancia de *L. crispatus* y

L. iners, respectivamente; sin embargo, nuestro CST-II estaba compuesto por una mezcla de varias especies de *Lactobacillus* y otros comensales (Figura 28). Además, otros cuatro CSTs no estaban dominados por especies de *Lactobacillus*: CST-IV (*Escherichia. coli*), CST-V (*Gardnerella vaginalis*), CST-VI (*Lactococcus lactis*) y CST-VII (*Klebsiella aerogenes*) (Figura 28). Al relacionar los CSTs con los resultados reproductivos, se observó que los CSTs-III y -V se asociaban con mayor frecuencia a mujeres con fracaso reproductivo (Figura 29), mientras que las mujeres con una microbiota endometrial de tipo CST-I, y -II, caracterizada por la presencia *L. crispatus* o de varios *Lactobacillus* junto con otros comensales, presentaban mejores tasas de recién nacido vivo. La abundancia general de *Lactobacillus* spp. ha sido propuesta como un marcador para el éxito reproductivo tanto en microbiomas vaginales como endometriales (Moreno et al., 2022; Petrova et al., 2015; Zeng et al., 2022). Varios estudios han vinculado la presencia de *L. crispatus*, el taxón dominante en CST-I, con un estado saludable en mujeres en edad reproductiva basado en su capacidad de prevenir infecciones de transmisión sexual y vaginosis bacteriana, así como en una asociación con una menor incidencia de parto prematuro (Fettweis et al., 2019; Golob et al., 2023; Ravel et al., 2021). A diferencia de otras especies del género, el papel de *L. iners* en el tracto genital femenino sigue en debate, dado que se ha descrito a esta bacteria como un patógeno oportunista y en la transición entre CSTs dominados por *Lactobacillus* y no dominados por *Lactobacillus* (Gajer et al., 2012, N. Zheng et al., 2021). *G. vaginalis*, la bacteria dominante en el CST-V, es uno de los patógenos más comunes en el tracto reproductivo femenino que se asocia con vaginosis bacteriana (Swidsinski et al., 2013). Cabe destacar que varios estudios han asociado la presencia de *G. vaginalis* en el endometrio con fracaso reproductivo (Garcia-Grau et al., 2019;

Liu et al., 2021; Lozano et al., 2023; Moreno et al., 2022) y adenomiosis (Kunaseeth et al., 2022), lo que coincide con nuestros resultados.

5.2 Análisis funcional: microbioma

Cuando ampliamos nuestro análisis e incluimos datos funcionales a partir de los genomas bacterianos, observamos un mayor número de asociaciones entre taxones y funciones en los tres grupos de fracaso reproductivo (NP, BP y CM) en comparación con el éxito reproductivo (LB), lo que sugiere que la disbiosis relacionada con la infertilidad puede desencadenar un desequilibrio en la señalización molecular no solo entre microbios, sino también con el huésped (Figuras 38-41). En este sentido, los transportadores ABC (*ATP-binding cassette*) mostraron correlaciones negativas con comensales (como *Cutibacterium acnes*, *Lactobacillus gasseri* y *Lactobacillus hominis*) en BP, pero asociaciones positivas con bacterias disbióticas (por ejemplo, *Bifidobacterium* spp. y *G. vaginalis*) en CM (Figuras 40 y 41). Es importante destacar que estos transportadores tienen una amplia gama de funciones, como la captación de micronutrientes y la promoción de resistencia a antibióticos mediante el eflujo de fármacos (Locher, 2016; Orelle et al., 2019). Por lo tanto, estos transportadores podrían tener un impacto significativo en la capacidad de las comunidades bacterianas para habitar un ecosistema. Por último, asociaciones positivas de transposasas con *Lactobacillus* spp. en diferentes resultados reproductivos (Figuras 38-40) sugieren que estos genes juegan un papel importante en el equilibrio de la composición de la microbiota endometrial. Estas transposasas son elementos genéticos móviles que tienen un papel importante en la plasticidad del genoma bacteriano y la adaptación al huésped (Aziz et al., 2010), y podrían ser beneficiosas para que los *Lactobacillus* establezcan una colonización en el entorno endometrial.

En cuanto a los factores de virulencia, hemos observado 360 genes significativamente enriquecidos en pacientes que sufren de fracaso reproductivo, además de 800 genes únicos en NP, BP y CM que no estaban presentes en muestras de pacientes que lograron un recién nacido vivo (Tablas 13, 15 y 16). Entre los diversos genes que mostraron abundancia diferencial entre los grupos, aquellos factores asociados con la adherencia, la modulación inmune, sistemas de transporte de moléculas y la formación de biofilms son algunos de los más representativos (Tablas 13-16 y Anexo 4).

Por ejemplo, ya se han asociado los biofilms con la etiología de la vaginosis bacteriana, ya que esta se asocia con la presencia de biofilms polimicrobianos en la vagina y el endometrio de las mujeres afectadas (Jung et al., 2017; Swidsinski et al., 2013). Aunque el impacto de las biopelículas bacterianas en la cavidad uterina aún no se ha caracterizado, se hipotetiza que podrían actuar como una barrera física contra la implantación del embrión, afectando negativamente el establecimiento del embarazo y los resultados reproductivos (Swidsinski et al., 2013). Además, el sistema de entrega de moléculas manipula los procesos de las células huésped para promover la supervivencia y proliferación bacteriana, permitiendo a los patógenos ejercer un control preciso sobre las funciones de las células huésped (Coulthurst, 2019; Schindler et al., 2023; Wagner et al., 2018). Así, su mayor abundancia entre los pacientes del grupo con fallo reproductivo sugiere que su modulación de las células huésped podría impactar significativamente los resultados reproductivos.

También encontramos una mayor abundancia de 2 genes bacterianos relacionados con la resistencia a antibióticos en pacientes que sufren de fracaso reproductivo, además de 57 genes que solo fueron detectados en este grupo en comparación con mujeres con LB (Tablas 17-21). Dada la creciente cantidad de datos

sobre la influencia del microbioma endometrial en la medicina reproductiva y el uso actual de antibióticos durante el tratamiento de fertilidad, estas características podrían ser de gran importancia para la práctica clínica.

En el futuro, estos resultados podrían dar lugar a nuevos métodos de diagnóstico o incluso a terapias dirigidas contra proteínas bacterianas específicas que se asocien significativamente con el fracaso reproductivo. No obstante, cabe señalar que la abundancia de genes no siempre es indicativa de la actividad transcripcional del microbioma, y los datos de metatranscriptómica serían una herramienta útil para evaluar la funcionalidad del microbioma endometrial y su relación con las células y el ambiente uterino.

5.3 Transcriptómica en el endometrio disbiótico

Se ha descrito que el microbioma puede modular varios mecanismos del huésped relacionados con la organización de la cromatina, la metilación del ADN y la expresión de proteínas para crear un nicho más beneficioso para la infección (Nichols & Davenport, 2021), lo que ya se ha asociado con la fisiopatología de diferentes tipos de cáncer (Guo et al., 2023; Park et al., 2022). En nuestro subconjunto de pacientes, hemos encontrado 90 genes que se expresaron diferencialmente en pacientes que sufrieron un fracaso reproductivo en comparación con aquellos que tuvieron un nacimiento vivo (Tablas 22-24). De estos, los más representativos son los genes asociados a la respuesta inmune, incluyendo la inmunidad humoral y celular, la secreción de citocinas y quimiocinas, la fagocitosis y respuesta de defensa a bacterias. Se ha demostrado que alteraciones en el microbioma del tracto reproductivo pueden modular factores proinflamatorios, lo que lleva a una alteración en la barrera epitelial y altos niveles de inflamación (Shannon et al., 2017; Wang et

al., 2021). Además, la disbiosis vaginal se ha asociado significativamente a la expresión de citocinas proinflamatorias en mujeres que padecen parto prematuro (Fettweis et al., 2019). De acuerdo con esto, estudios *in vitro* con un modelo de colonización endometrial demostró que el co-cultivo de *Lactobacillus rhamnosus* con células epiteliales endometriales modula significativamente los niveles de IL-6, IL-8 y otras interleucinas, lo que sugiere su uso factible como probiótico en la salud ginecológica (Chenoll et al., 2019).

Además, el grupo de pacientes que sufre de fracaso reproductivo también mostró una expresión diferencial de múltiples genes codificantes de GPCRs (Tablas 22-24). Evidencias científicas han demostrado la relación de GPCRs con la viabilidad de células endometriales, el ciclo menstrual y la decidualización (Aikawa et al., 2017; Sprague et al., 2021; Totorikaguena et al., 2017), pero también su relación con enfermedades ginecológicas u obstétricas (Aghajanova et al., 2016; Xu et al., 2021; Yin et al., 2022). Además, también se ha propuesto la inhibición de GPCRs específicos como un tratamiento factible para estas enfermedades (Conrad, 2016; Tapmeier et al., 2021). Así, alteraciones de la señalización de GPCRs en nuestro subconjunto de pacientes coincide con la literatura previa y añade más evidencia de que dirigirse a estas vías de señalización podría tener un importante impacto clínico en medicina reproductiva.

5.4 Fortalezas

- La mayoría de los artículos publicados que se centran en el microbioma emplean la secuenciación del gen ARNr 16S que ofrece únicamente información taxonómica; sin embargo, nuestro estudio evaluó la composición de la comunidad

bacteriana utilizando secuenciación de metagenomas que permite la evaluación funcional del microbioma endometrial.

- Nuestro trabajo incluye el uso de controles negativos, positivos y blancos de extracción recomendados en las buenas prácticas para el análisis del microbioma endometrial (Molina et al., 2021).
- Gran cobertura de secuenciación, con aproximadamente 150 millones de lecturas por muestra en análisis de metagenomas, lo que posibilita el análisis del microbioma uterino, conocido por ser de baja biomasa; y 100 millones de lecturas por muestra en análisis de metatranscriptómica, facilitando el análisis del tejido uterino.
- La toma de muestras endometriales y la transferencia embrionaria en nuestro estudio se realizó en el momento de máxima receptividad endometrial, lo que mitiga posibles sesgos transcripcionales asociados al momento de la toma de muestras.

5.5 Limitaciones

- Tamaño muestral limitado en los grupos de BP y CM, lo que dificultó la evaluación de diferencias significativas en la composición y función del microbioma entre las muestras de estos resultados reproductivos.
- No se puede descartar completamente la contaminación vaginal al evaluar el microbioma endometrial a través de la recolección transcervical
- No todos los embriones transferidos habían sido evaluados cromosómicamente, lo que podría ser un factor de confusión en la evaluación de los resultados reproductivos en nuestro estudio.

- El análisis utilizando Diamond/BLAST es altamente sensible, pero no permite asignar genes bacterianos a los taxones de los que provienen, por lo que solo pudimos inferir la composición genética de las comunidades completas.
- Debido a la baja biomasa presente en la microbiota endometrial, y la alta presencia de ARN ribosomal, el número de lecturas bacterianas en el análisis de metatranscriptómica no fue suficiente para una cobertura completa de los transcriptomas bacterianos, por lo que se presentan únicamente los resultados de expresión génica del endometrio humano.

Los resultados de este trabajo contribuyen al conocimiento de una huella molecular en la que la detección/abundancia de genes bacterianos específicos podría utilizarse como biomarcadores de los resultados reproductivos. Nuestros datos también sugieren posibles disfunciones celulares en el endometrio que podrían estar asociadas al microbioma presente en la cavidad uterina de pacientes que sufren fracaso reproductivo. Estos resultados ofrecen nuevas perspectivas sobre la fisiopatología de la infertilidad, desentrañando nuevos mecanismos moleculares con impacto en reproducción humana y abriendo camino para nuevos enfoques de tratamiento.

6. Conclusiones

- Los resultados del estudio taxonómico mediante metagenómica confirman la existencia de una microbiota endometrial diferencial, a nivel de especie, en pacientes con fracaso reproductivo caracterizado por un enriquecimiento significativo de *G. vaginalis*, *B. breve*, *B. bifidum* y *A. vaginae*, entre otras bacterias disbióticas.

- La composición de CSTs en el endometrio muestra cuatro grandes comunidades principalmente dominadas por *L. crispatus*, *L. iners*, *G. vaginalis*, o una mezcla de distintos comensales. Las CSTs dominadas por *L. iners* y *G. vaginalis* están más asociadas a fracaso reproductivo que *L. crispatus* y otras comunidades comensales.
- Los estudios funcionales de metagenómica destacan la importancia de considerar la comunidad bacteriana en su totalidad ya que una misma bacteria puede asociarse de manera positiva (coexistencia) o negativa (exclusión) con otras bacterias conduciendo a perfiles genómicos diferenciales en distintos resultados reproductivos.
- En pacientes con fracaso reproductivo se observa un enriquecimiento significativo de genes asociados a la resistencia a antibióticos, principalmente los asociados a bombas de eflujo de antibióticos mediante transportadores ABC, MFS (*Major Facilitator Superfamily*) y RND (*resistance-nodulation-cell division*).
- Los principales factores de virulencia significativamente enriquecidos en fracaso reproductivo afectan a la modulación del sistema inmune, adherencia y formación de *biofilms*, factores nutricionales, y sistemas de entrega de metabolitos y moléculas.
- En el endometrio de pacientes con fracaso reproductivo existe una amplia desregulación de genes implicados en el sistema inmune, incluyendo la inmunidad humoral y celular, la secreción de citocinas y quimiocinas, la fagocitosis y respuesta de defensa a bacterias.
- En mujeres con fracaso reproductivo existen diferencias en la expresión de GPCRs (incluyendo receptores de ácidos grasos, angiotensina, glucagón y ácido hidrocarboxílico) en las células endometriales.

INDEX

1. INTRODUCTION.....	43
1.1 THE HUMAN UTERUS	43
1.1.1 Endometrial receptivity	46
1.2.1 The microbiome in human health.....	52
1.2.2 Technical approaches for microbiome assessment.....	53
1.2.3 Challenges of studying the endometrial microbiota	60
1.2.3.1. <i>Population and Study Design</i>	60
1.2.3.2. <i>Sample Handling and Experimental Design</i>	61
1.3 THE FEMALE REPRODUCTIVE TRACT MICROBIOTA.....	64
1.3.1 Composition of the healthy endometrial microbiota	68
1.3.2 Implications of endometrial microbiota in human infertility	71
1.3.2.1. <i>Endometrial microbiota and reproductive outcomes</i>	73
1.3.3 Implications of endometrial microbiota in gynecological conditions	76
1.3.4 Tools for endometrial dysbiosis management.....	79
1.4 GENETIC AND FUNCTIONAL DIVERSITY WITHIN THE MICROBIOME.....	83
1.4.1 Antibiotic resistance	84
1.4.2 Virulence factors	86
1.4.3 Microbial factors and host cell modulation	92
1.4.4 Host cell response to dysbiotic microbiota and infection.....	98
1.4.5 Applications in reproductive medicine.....	101
2. BACKGROUND AND HYPOTHESIS.....	102
2.1 BACKGROUND	102
2.2 HYPOTHESIS.....	102
3. OBJECTIVES.....	103
4. METHODS.....	104

4.1 PATIENTS, SAMPLES, AND ETHICS	104
4.2 GENOMIC DNA AND RNA ISOLATION	109
4.3 WHOLE METAGENOME SEQUENCING	110
4.4 METATRANSCRIPTOMICS	110
4.5 BIOINFORMATIC SEQUENCING ANALYSIS: METAGENOMICS.....	111
4.5.1 Sequencing-based microbiome analysis	111
4.5.2 Community state types (CSTs)	112
4.5.3 Correlation network analysis.....	112
4.5.4 Virulence factors and antibiotic resistance analysis.....	113
4.6 BIOINFORMATIC SEQUENCING ANALYSIS: METATRANSCRIPTOMICS	113
4.6.1 Sequencing-based RNA analysis	113
4.7 STATISTICAL ANALYSIS.....	114
5. RESULTS	115
5.1 CHARACTERIZATION OF ENDOMETRIAL BACTERIAL COMMUNITIES..	115
5.1.1 Bacterial species associated with reproductive outcomes	122
5.1.2 Taxonomic networks and microbe coexistence.....	123
5.2 GENETIC AND FUNCTIONAL DIVERSITY OF BACTERIAL COMMUNITIES IN REPRODUCTIVE OUTCOMES	132
5.2.1 Bacterial physiological functions	132
5.2.2 Functional profiling of bacterial taxa	142
5.2.3 Virulence factors	146
5.2.4 Antibiotic resistance	153
5.3 TRANSCRIPTOMICS IN THE DYSBIOTIC ENDOMETRIUM	157
6. DISCUSSION	167
7. CONCLUSIONS	185
8. REFERENCES	187
9. ANNEXES	232

ABBREVIATIONS

	16S rRNA	16S ribosomal RNA
	β-hCG	Beta human chorionic gonadotropin
A	ABC	ATP-binding cassette
	ADRB2	Adrenergic β2 receptor
	AMP	Antimicrobial peptide
	AMR	Antimicrobial resistance
	APC	Antigen-presenting cell
	ARG	Antibiotic resistance gene
	ARO	Antibiotic resistance ontology
	ART	Assisted reproductive technologies
B	BMI	Body mass index
	BP	Biochemical pregnancy
	BV	Bacterial vaginosis
	BWA	Burrows-Wheeler Aligner
C	CARD	Comprehensive antibiotic resistance database
	CASR	Calcium-sensing receptor
	cDNA	Complementary DNA
	CE	Chronic endometritis
	CM	Clinical miscarriage
	COVID-19	Coronavirus disease 2019
	CP	Chloramphenicol
	CRLR	Calcitonin receptor-like receptor
	CST	Community state type
	CV	Cervical mucus drawn from the cervical canal
D	DNA	Deoxyribonucleic acid
E	E ₂	Estrogen
	EB	Endometrial biopsy
	EF	Endometrial fluid
	EMMA	Endometrial microbiome metagenomic analysis
	ER	Estrogen receptor
	ERA	Endometrial Receptivity Analysis
	ESHRE	European Society of Human Reproduction and Embryology
	ET	Embryo transfer

F	FDR	False discovery rate
	FET	Frozen embryo transfer
	FFT	Fecal filtrate transplant
	FGT	Female genital tract
	FLASH	Fast Length Adjustment of Short reads
	FMT	Fecal microbiota transplant
G	GO	Gene ontology
	GPCR	G protein-coupled receptor
H	HCG	Human chorionic gonadotrophin
	HK	Histidine kinase
	HMP	Human Microbiome Project
	HRT	Hormone replacement therapy
I	ICSI	Intracytoplasmatic sperm injection
	IFN	Interferon
	iHMP	Integrative human microbiome project
	IL	Interleukin
	IVF	<i>In vitro</i> fertilization
K	KEGG	Kyoto encyclopedia of genes and genomes
L	LB	Live birth
	LBR	Live birth rate
	LD	<i>Lactobacillus</i> -dominated
	LGT	Lower genital tract
	LIF	Leukemia inhibitory factor
	LPAR3	Lysophosphatidic acid receptor
M	MII	Metaphase II
	miRNA	Micro RNA
	MFS	Major facilitator superfamily
	MOR	Mu opioid receptor
	MTX	Metatranscriptomics
	MZ	Metronidazole
N	NA	Not applicable
	NGA	Next generation microbial
	NGS	Next generation sequencing
	NIH	National Institutes of Health
	NK	Natural killer

	NK1R	Tachykinin receptor 1
	NLD	Non- <i>Lactobacillus</i> -dominated
	NP	No pregnancy
	NPSR1	Neuropeptide S receptor 1
O	OMV	Outer membrane vesicle
	OTU	Operational taxonomic unit
P	P4	Progesterone
	pET	Personalized embryo transfer
	PF	Peritoneal fluid
	PID	Pelvic inflammatory disease
	PKR1	Prokineticin 2 receptor
	PR	Progesterone receptor
	PRR	Pattern recognition receptors
Q	QS	Quorum sensing
R	RIF	Repeated implantation failure
	RGI	Resistance Gene Identifier
	RNA	Ribonucleic acid
	RND	Resistance-nodulation-division
	RPL	Recurrent pregnancy loss
	rRNA	Ribosomal ribonucleic acid
	RXFP1	Relaxin/insulin-like family peptide receptor 1
S	SCFA	Short-chain fatty acid
	STI	Sexually transmitted infection
T	TCS	Two-component system
	TFF3	Trefoil factor 3
	TGF- β	Transforming growth factor- β
	TLR	Toll-like receptor
U	UGT	Upper genital tract
V	VFDB	Virulence factor database
	VMT	Vaginal microbiota transplantation
W	WMS	Whole metagenome sequencing
	WOI	Window of implantation

LIST OF FIGURES

Figure 1. Schematic representation of the female genital tract.	43
Figure 2. Temporal transcriptome dynamics of endometrial transformation across the human menstrual cycle.	45
Figure 3. Processes and molecules implicated in the acquisition of endometrial receptivity.....	47
Figure 4. Human microbiota composition in different locations.	50
Figure 5. Factors influencing microbiome composition.....	51
Figure 6. Variable regions of the 16S ribosomal RNA.....	54
Figure 7. Comparison between different techniques for microbiome analysis.	55
Figure 8. Major components of whole-metagenome sequencing studies.	56
Figure 9. Summary of metatranscriptomics workflow.	59
Figure 10. Sources of contamination in endometrial microbiota	62
Figure 11. Microbiota composition of the female genital tract.....	65
Figure 12. Bacterial composition of the thirteen sub-CSTs described in the vaginal microbiota.....	67
Figure 13. Heatmap depicting endometrial microbiota distribution during the acquisition of endometrial receptivity.....	70
Figure 14. Overview of bacterial genera associated with different reproductive outcomes.	75
Figure 15. <i>Fusobacterium nucleatum</i> and its association with endometriotic lesions in a mouse model	78
Figure 16. Examples of antibiotic resistance mechanisms.	85
Figure 17. Examples of virulence factors in a bacterial cell.....	87
Figure 18. Model of biofilm formation.....	88
Figure 19. Different cellular and extracellular targets of next generation antimicrobials (NGAs) and the impact on different virulent phenotypes.	92
Figure 20. Potential mechanisms underlying host gene expression–microbiome associations.	93

Figure 21. Potential mechanisms linking endometrial microbiota to early pregnancy loss	95
Figure 22. Diversity of G-protein-coupled receptor signaling.....	96
Figure 23. Overview of some microbial-derived metabolites known to interact with human receptors	97
Figure 24. Scheme of host mechanisms for microbiota maintenance and regulation in the gut	99
Figure 25. Patients allocated in the study classified according to their reproductive outcomes after assisted reproductive technology.....	106
Figure 26. Patients allocated in each reproductive outcome group after library preparation.	116
Figure 27. Distribution of CSTs in endometrial samples from patients with distinct reproductive outcomes..	117
Figure 28. Taxonomic composition of the samples in each community state type.....	119
Figure 29. Differential presence of CSTs between reproductive success and failure..	120
Figure 30. Taxonomic networks within samples from the LB group.	124
Figure 31. Microbe coexistence within samples from the LB group.	125
Figure 32. Taxonomic networks within samples from the NP group.	126
Figure 33. Microbe coexistence within samples from the NP group.	127
Figure 34. Taxonomic networks within samples from the BP group.	128
Figure 35. Microbe coexistence within samples from the BP group.	129
Figure 36. Taxonomic networks within samples from the CM group.	130
Figure 37. Microbe coexistence within samples from the CM group.	131
Figure 38. Coupled taxonomic and functional networks for the bacterial communities in the endometrium of infertile women from the LB group.	143
Figure 39. Coupled taxonomic and functional networks for the bacterial communities in the endometrium of infertile women from the NP group.	144
Figure 40. Coupled taxonomic and functional networks for the bacterial communities in the endometrium of infertile women from the BP group.	145

Figure 41. Coupled taxonomic and functional networks for the bacterial communities in the endometrium of infertile women from the CM group.....	146
Figure 42. Virulence factors genes enriched in NP compared to LB.....	148
Figure 43. Virulence factors genes enriched in CM compared to LB.....	151
Figure 44. Differential gene expression of endometrial samples from the NP group...	159
Figure 45. Differential gene expression of endometrial samples from the BP group...	163
Figure 46. Differential gene expression of endometrial samples from the CM group..	165
Figure 47. Overview of the most important findings from women who suffer from reproductive failure.....	184

LIST OF TABLES

Table 1. Pregnancy outcomes after frozen embryo transfer in patients with or without personalized treatment recommendations based on the microbial 16S rRNA gene sequencing	81
Table 2. Overview of GPCRs previously associated with gynecological processes and diseases.	98
Table 3. Sociodemographic and clinical characteristics of study participants.....	106
Table 4. Clinical variables in patients with different reproductive outcomes.....	108
Table 5. <i>G. vaginalis</i> abundance in endometrial samples.	121
Table 6. Taxa significantly enriched in reproductive failure compared to reproductive success.	122
Table 7. Bacterial genes enriched in the live birth group compared to the no pregnancy group using DeSeq2 analysis.	133
Table 8. LEfSe analysis describing bacterial genes enriched in each reproductive outcome.	134
Table 9. Bacterial functions and associated genes significantly enriched in LB.....	136
Table 10. Bacterial functions and associated genes significantly enriched in NP.	137
Table 11. Bacterial functions and associated genes significantly enriched in BP.	138
Table 12. Bacterial functions and associated genes significantly enriched in CM.	140
Table 13. Number of virulence factor-associated genes in the NP group compared to LB.	147
Table 14. Virulence factors differentially abundant in NP compared to LB.	148
Table 15. Number of virulence factor-associated genes in the BP group.....	150
Table 16. Number of virulence factor-associated genes in the CM group.	150
Table 17. Antibiotic resistance-related genes differentially abundant in NP compared to LB.....	153
Table 18. Antibiotic resistance-related genes present only in samples from NP group compared to LB.	154

Table 19. Antibiotic resistance-related genes present only in samples from the BP group compared to LB.....	155
Table 20. Antibiotic resistance-related genes differentially abundant in CM compared to LB.....	155
Table 21. Antibiotic resistance-related genes present only in samples from the CM group compared to LB.....	156
Table 22. Biological processes significantly up- and downregulated in no pregnancy compared to live birth.....	160
Table 23. Biological processes significantly upregulated in biochemical pregnancy compared to live birth.....	164
Table 24. Biological processes significantly upregulated in clinical miscarriage compared to live birth.....	166

1. INTRODUCTION

1.1 THE HUMAN UTERUS

The uterus, a muscular and hollow organ situated in the upper female genital tract, serves as a pivotal site for various reproductive processes, including menstruation, implantation, pregnancy, and labor. It is composed of three distinct layers, each contributing to its structural integrity and physiological function (Moreno et al., 2023) (Figure 1).

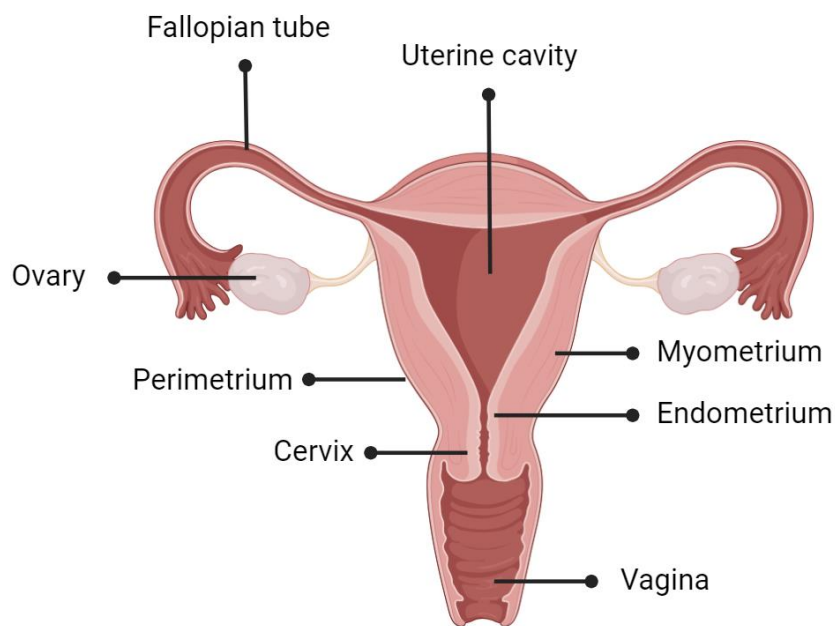


Figure 1. Schematic representation of the female genital tract. Created with BioRender.com.

The perimetrium, the outermost layer, is continuous with the peritoneal cavity, providing protection and support to the uterine structure. Beneath the perimetrium lies the myometrium, a layer of smooth muscle that undergoes significant changes throughout pregnancy to accommodate fetal growth and facilitate labor, being

responsible for uterine contractions during childbirth. Lastly, the endometrium, lining the interior of the uterine cavity, which is divided in two stratum: the *stratum functionalis*, a highly vascularized mucosal layer that is shed during menstruation along the woman's reproductive age, from menarchy to menopause, and the *stratum basalis*, adjacent to the myometrium, which is retained and regenerates the functional layer after menstruation (Critchley et al., 2020).

The endometrial tissue – comprising luminal and glandular epithelium, stromal fibroblasts, vascular smooth muscle and endothelium, stem/progenitor cells, and immune cells – undergoes dynamic changes every month. Historically, the menstrual cycle has been divided into proliferative and secretory phases. During the proliferative phase, estrogen promotes stromal cell proliferation and elongation of spiral arteries, whereas in epithelial cells this is associated with proliferation, differentiation, and cellular apoptosis. The early secretory phase is marked by a rise in progesterone levels, as well as increased cell metabolism and negative cell proliferation regulation. On the other hand, the mid-secretory phase involves significant metabolic secretory and non-proliferative activity, mainly associated with innate immune response and the response to stress and wounding; while the late secretory phase is marked by a decrease in estrogen and progesterone levels, leading to extracellular matrix degradation, an inflammatory response, and apoptosis, culminating in menstruation (Moreno et al., 2023).

Although embryo quality is crucial for the achievement of a live birth, there is increasing knowledge about the impact of endometrial physiology in reproductive health. In fact, throughout each menstrual cycle, dynamic molecular changes occur within the endometrial tissue, which govern processes such as endometrial receptivity

and decidualization (W. Wang et al., 2020), and have a great impact on embryo implantation and successful pregnancy outcomes (Figure 2).

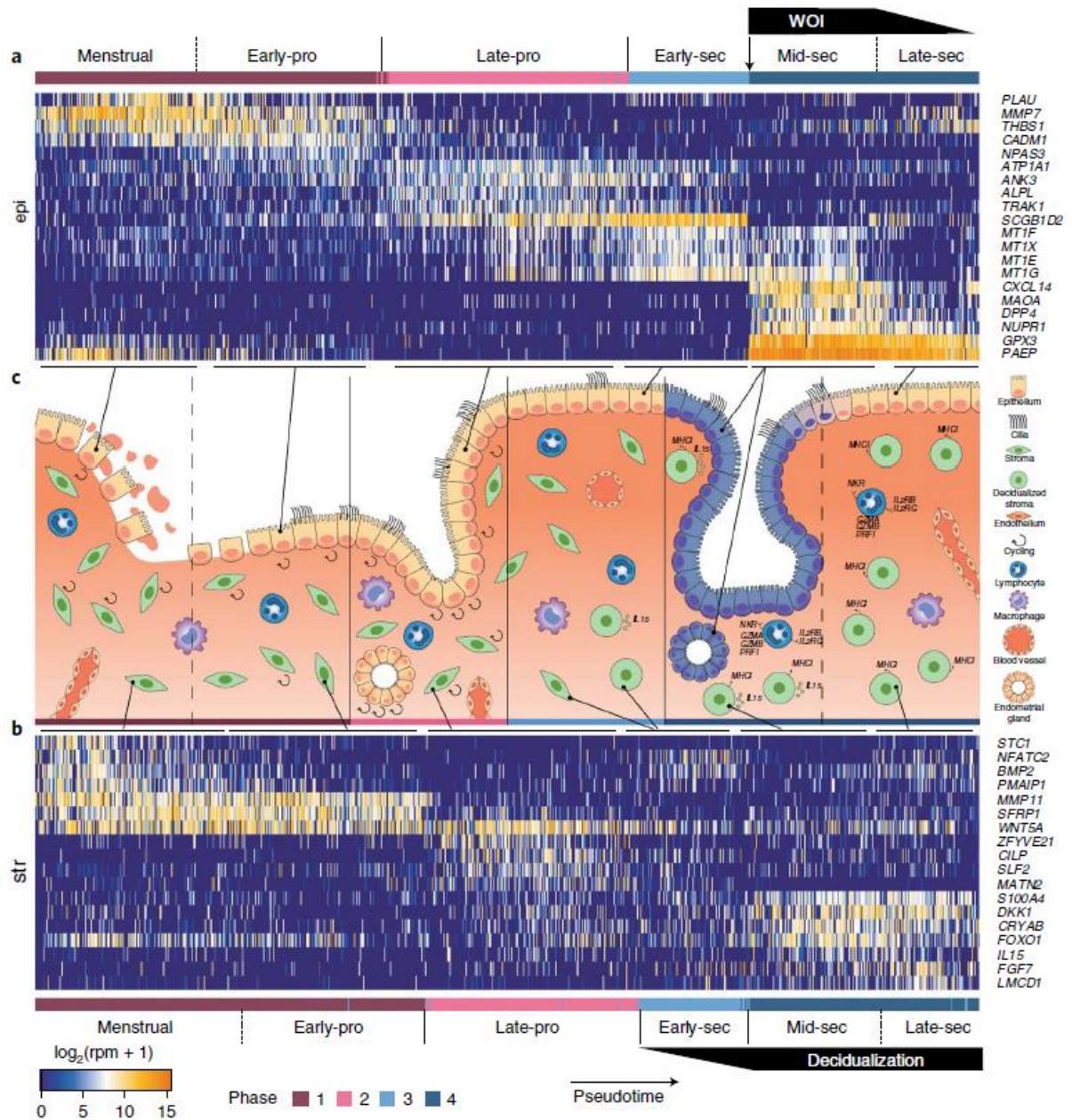


Figure 2. Temporal transcriptome dynamics of endometrial transformation across the human menstrual cycle. Exemplary phase- and subphase-defining genes, and the relationship between single-cell transcriptomically defined and histologically defined endometrial phases for unciliated epithelia (A) and stromal fibroblasts (B) across a human menstrual cycle (C). Reprinted from Wang et al., 2020 with permission from Springer Nature.

1.1.1 Endometrial receptivity

Throughout the menstrual cycle, hormonal changes orchestrate the activity of several molecular pathways within the female genital tract (FGT). Of particular significance is a narrow timeframe, called the window of implantation (WOI), that extends from three to six days. At this critical period, the endometrium reaches its peak of receptivity, facilitating blastocyst adherence and trophoblast invasion, ultimately enabling the establishment of an early healthy pregnancy. Importantly, certain conditions may displace or narrow this window, which can significantly impact human fertility. Thus, the study of the molecular dynamics driving uterine physiology stand crucial for the improvement of reproductive outcomes (Lessey & Young, 2019).

There are several morphological and molecular processes that facilitate the highly complex phenomenon of pregnancy (Figure 3). Among them, there is a differential expression of adhesion molecules such as cadherins and integrins, which may be tightly related to the first attachment of the embryo to the epithelial endometrium (Moreno et al., 2023). Nonetheless, the expression of antiadhesion molecules also fluctuates during the menstrual cycle, and research has demonstrated a lower expression of MUC1 expression in the glandular and luminal epithelium in women who experience repeated implantation failure (RIF) when compared to women with recurrent miscarriage, suggesting that MUC1 is associated with blastocyst selection by the endometrium (F. Wu et al., 2018).

Another essential process involves chemokines, cytokines, and growth factors abundance (Figure 3). For instance, cytokines are released by the maternal tissues and produce paracrine signals that influence embryo development and maintenance of pregnancy, and dysregulation of cytokine production could have negative impacts on human conception (Moreno et al., 2023). Additionally, altered expression of

interleukins such as IL-6 and IL-11 has been observed in women suffering from recurrent miscarriage (Linjawi et al., 2004; Wolff, 2000), while high levels of IL-4, IL-5, and IL-10 seem to be crucial in the maintenance of a healthy pregnancy (Piccinni et al., 2021). Conversely, some growth factors exhibit changes in their expression patterns around the implantation period, suggesting they play important functions at this process (Moreno et al., 2023). For instance, transforming growth factor- β (TGF- β) signaling promotes stromal decidualization (Ni & Li, 2017); and its decreased expression has been demonstrated in patients with RIF (F. Guo et al., 2018).

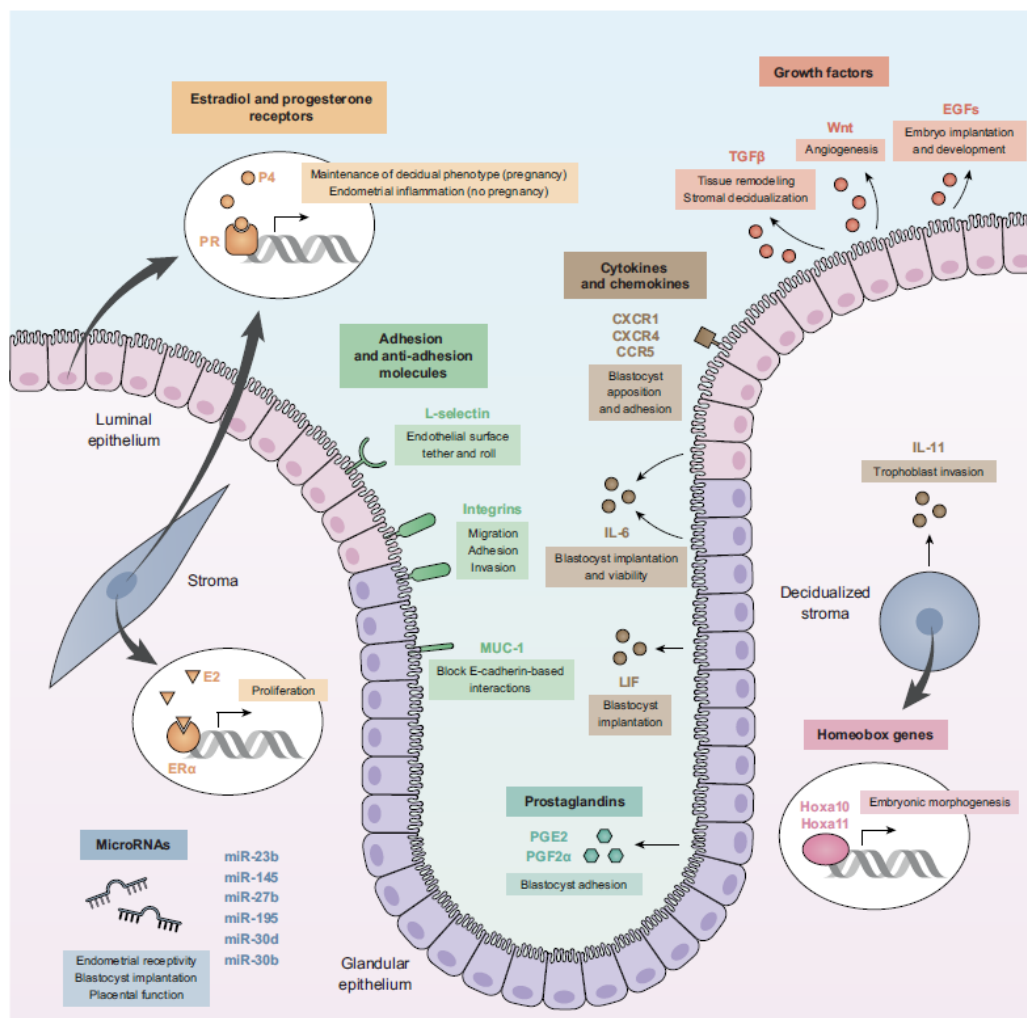


Figure 3. Processes and molecules implicated in the acquisition of endometrial receptivity. Abbreviations: E₂, estrogen; ER, estrogen receptor; LIF, leukemia inhibitory factor; P₄, progesterone; PR, progesterone receptor; TGF- β , transforming growth factor- β . Reprinted from Moreno et al., 2023 with permission from The American Physiological Society and Copyright Clearance Center.

The expression of estradiol and progesterone receptors within the uterine tissues also play a role on pregnancy (Figure 3). E₂ directs early and late proliferative phases in the endometrium, acting on several levels such as cell cycle gene transcription, DNA synthesis, cell proliferation, and immune cell infiltration. Progesterone receptors, on the other hand, are crucial for endometrial maturation and maintenance of the decidual phenotype of stromal cells during the late luteal phase, facilitating the development of a healthy pregnancy (Moreno et al., 2023). Moreover, it has been demonstrated that the expression of epithelial progesterone receptors decreases during the days before implantation; but increases in the endometrial stroma surrounding the implantation site in a murine model (Takamoto et al., 2002).

The phenomenon of endometrial receptivity also includes the remodeling of the endometrial stroma, which differentiates into decidualized stromal cells, and involves morphological and molecular shifts to facilitate successful embryo implantation (Figure 3). The decidua is a complex tissue that is closely connected to the fetoplacental region, and is crucial for trophoblast invasion and maintenance of a healthy pregnancy. Conversely, defective decidualization has been associated with different reproductive and obstetric disorders (Ng et al., 2020). Interestingly, proteomic analysis revealed distinct expression patterns of 60 proteins during decidualization, including the secretion of growth factors by endometrial glands and of canonical decidualization markers such as prolactin and insulin-like growth factor binding protein 1 (Garrido-Gomez et al., 2011); and defects on this process have been pointed out as potential contributors to the pathogenesis of pregnancy disorders such as severe preeclampsia (Garrido-Gómez et al., 2022).

Lastly, distinct expression of immune receptors and small molecules, alongside an infiltration of immune cells, has been described in the late secretory phase of the

menstrual cycle. For instance, differential expression of immune receptors such as KIR (in uterine cells) and HLA-C (in the trophoblast) seem to play a pivotal role in promoting fetal tolerance by the maternal immune system (Chazara et al., 2011). Moreover, an imbalance in Natural Killer (NK) cells numbers has been linked to excessive oxidative stress and reproductive failure in patients with recurrent reproductive failure (Quenby et al., 2008), and an infiltration of maternal macrophages was demonstrated to protect the growing embryo against their attack (Co et al., 2013). Considering the complex and intricate relation between the immune system and the maintenance of a healthy pregnancy, disruption in immune signaling may have a significant impact in reproductive failure. Building on this knowledge, it is feasible to speculate that changes in immune pathways due to bacterial colonization should also impact reproductive outcomes (Mor & Cardenas, 2010): although an optimal microbiota may serve as the primary line of defense against pathogens, an alteration in its composition might be associated with pathological states or conditions that impair organ physiology, embryo implantation, and pregnancy maintenance.

Taken together, all these changes are part of the preparation of the endometrial tissue for embryo attachment and establishment of a pregnancy. Thus, considering the endometrium undergoes natural fluctuations of up and downregulation of several genes throughout each cycle, it is important for studies in the field of reproductive medicine to establish a certain period for sample collection, mitigating heterogeneity and avoiding confounding factors. Nonetheless, in addition to chromosomal normalcy and endometrial receptivity, various factors such as immune tolerance, and the microbiota composition within the reproductive tract may significantly influence the success of implantation and pregnancy outcomes.

1.2 THE HUMAN MICROBIOME

The human microbiome can be defined as "the ecological community of commensal, symbiotic, and pathogenic microorganisms that literally share our body space" (Lederberg & McCray, 2001). Nowadays, the term microbiota characterizes the group of microorganisms that coexist in a defined environment, whilst microbiome defines the whole habitat, comprising microbes, their genomes, and the surrounding conditions (Marchesi & Ravel, 2015).

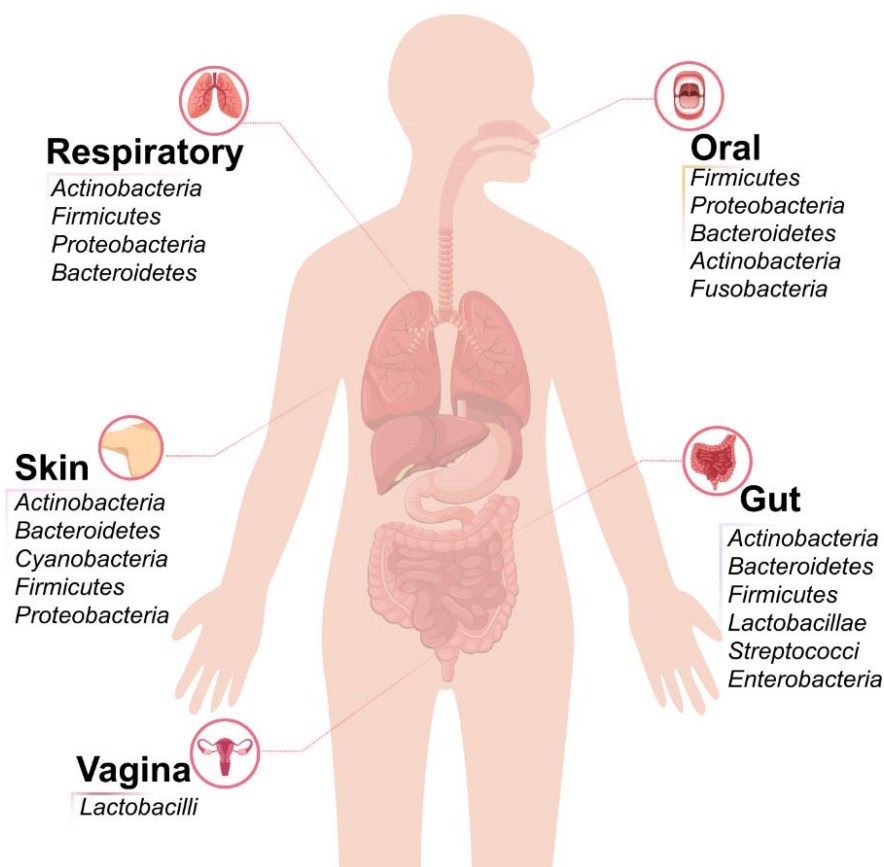


Figure 4. Human microbiota composition in different locations. Predominant bacteria in the oral cavity, respiratory tract, skin, gut, and vagina. Adapted from Hou et al., 2022, permission for figure reproduction granted by the Creative Commons license (<http://creativecommons.org/licenses/by/4.0/>).

The Human Microbiome Project (HMP), led by the National Institutes of Health (NIH) and launched in 2007, aimed to characterize the microbiome of five major body sites – the nasal passages, oral cavity, skin, gastrointestinal tract, and urogenital tract

- in healthy human subjects (NIH HMP Working Group et al., 2009). This project aimed to provide valuable data through the comprehensive analysis of global communities and raised the possibility of improving human health by monitoring/manipulating the human microbiome. Noteworthy, these communities include microorganisms such as fungi, archaea, bacteria, and viruses, but bacteria represent the most numerous and best-studied members of the human microbiota (Figure 4).

Since then, the presence of a core microbiota has been described in several body sites, such as the gut, oral cavity, female/male genital tract, and skin (El-Sayed et al., 2021a; Hou et al., 2022). Moreover, biological and environmental changes, such as aging, pH, geographic location, diet, use of antibiotics, and way of delivery may affect the composition of the highly dynamic human microbiome (Figure 5), which stands as a great challenge to study homogeneity and comparison of results (El-Sayed et al., 2021b).

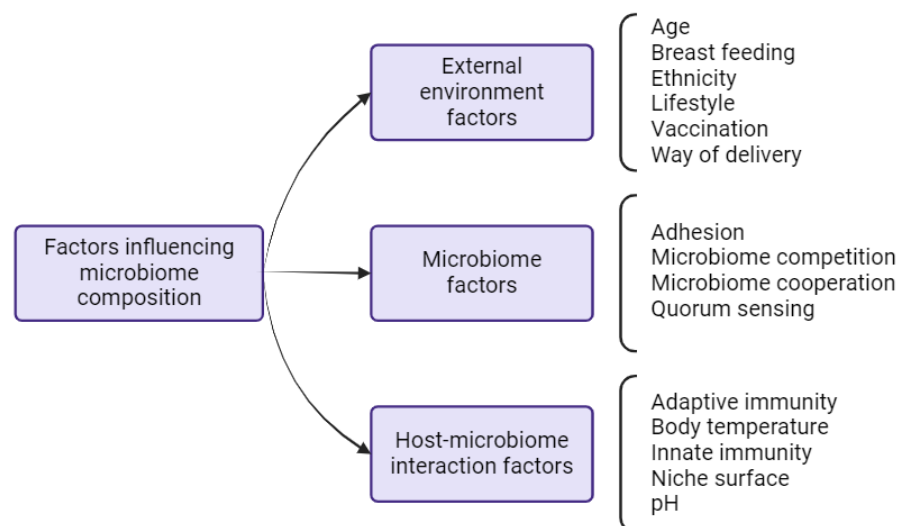


Figure 5. Factors influencing microbiome composition. Created with BioRender.com. Adapted from El-Sayed et al., 2021b with permission from Springer Nature.

1.2.1 The microbiome in human health

Commensal bacteria play fundamental roles in human health when in proper balance. For example, gut microbes play an important role in synthesis and absorption of essential vitamins and metabolism of nutrients, but also support immune system functions and prevents bacteria invasion by maintaining the intestinal epithelium integrity (Rinninella et al., 2019). Additionally, the skin microbiota acts as a barrier against invasion and infection by pathogenic microbes, promoting epithelial differentiation and creating a chemical environment that impairs foreign bacteria colonization (Harris-Tryon & Grice, 2022).

Given the important role of microbiome in human health, alterations in the composition of the microbiota or imbalances in microbial communities have been associated with the onset of numerous pathologies including asthma, coronavirus disease 2019 (COVID-19), atopic dermatitis, periodontitis, obesity, cancer, Crohn's disease, and autoimmune conditions (El-Sayed et al., 2021a; Manos, 2022). Curiously, the gut-brain axis, defined as a bidirectional communication network between the gastrointestinal tract and the central nervous system that plays a crucial role in metabolism, mood, behavior, and cognitive functions, has been demonstrated to be influenced by the microbiota. For instance, assessment of the microbiota-gut-brain axis revealed gut microbiota disbalance is implicated in complex behavior and neurological conditions such as Parkinson's disease, autism, anxiety, and Alzheimer's disease (Cryan et al., 2019). Thus, further understanding of the microbiota-gut-brain axis has led to the exploration of interventions such as probiotics and antibiotics to potentially improve mental health outcomes and overall brain physiology (Cryan et al., 2019).

Despite growing knowledge regarding the influence of microbiota composition in human health, little is known about the functional impact of microbial communities and genomes in disease onset. Building on the success of the HMP, the Integrative Human Microbiome Project (iHMP - launched in 2014) sought to support the further understanding of microbiome-host interactions using multi-omic technologies in three cohorts of microbiome-associated conditions: preterm birth, inflammatory bowel disease, and type 2 diabetes (The Integrative HMP (iHMP) Research Network Consortium, 2014). Thus, new advances in molecular biology and the use of multi-omic studies now focus on the mechanisms by which microbes may trigger human disease, and continuous research should provide valuable insight into new therapeutic approaches and prevention strategies.

1.2.2 Technical approaches for microbiome assessment

Methods to study the microbiota composition have evolved significantly throughout the years. Although microscopy based on gram-stained smears is a classical method for diagnosing bacterial vaginosis (BV) and other diseases, the information it provides may not always be sufficient and reliable (Savicheva, 2023). Additionally, microbiology became fully dependent on culture methods following the invention of plating techniques (Hiergeist et al., 2015). In fact, most of our knowledge about the human microbiota comes from culture-based approaches. Nevertheless, it has been estimated that between 20% and 60% of the human microbiome is uncultivable depending on the body site, and *in vitro* conditions may favor the growth of some strains over others (Hiergeist et al., 2015; NIH HMP Working Group et al., 2009). Thus, results provided by culture methods may bear an underestimation of the ecosystem diversity as well as failure in identification of microorganisms that may be associated

with a given disease. For this reason, although culture data is still informative and useful, molecular technologies such as next generation sequencing (NGS) are a milestone in microbiome-related studies, revealing an enormous amount of information that was not previously accessible due to method limitations (Wensel et al., 2022).

Sequencing of the 16S rRNA bacterial gene is the molecular technique mostly predominant in microbiota studies. This gene contains nine hypervariable regions (V1 to V9) located among highly conserved sequences that serve as a molecular fingerprint since they are species-specific (Figure 6) (Wensel et al., 2022). The resulting operational taxonomic units (OTUs) may be described as groups of similar variants of the 16S rRNA marker gene, each representing a taxonomic unit of a bacterium genus or species depending on the sequence similarity threshold: whilst a 97% identity threshold is typically required to distinguish bacteria at the genus level, species separation demands a higher threshold of 99% sequence identity (Edgar, 2018).

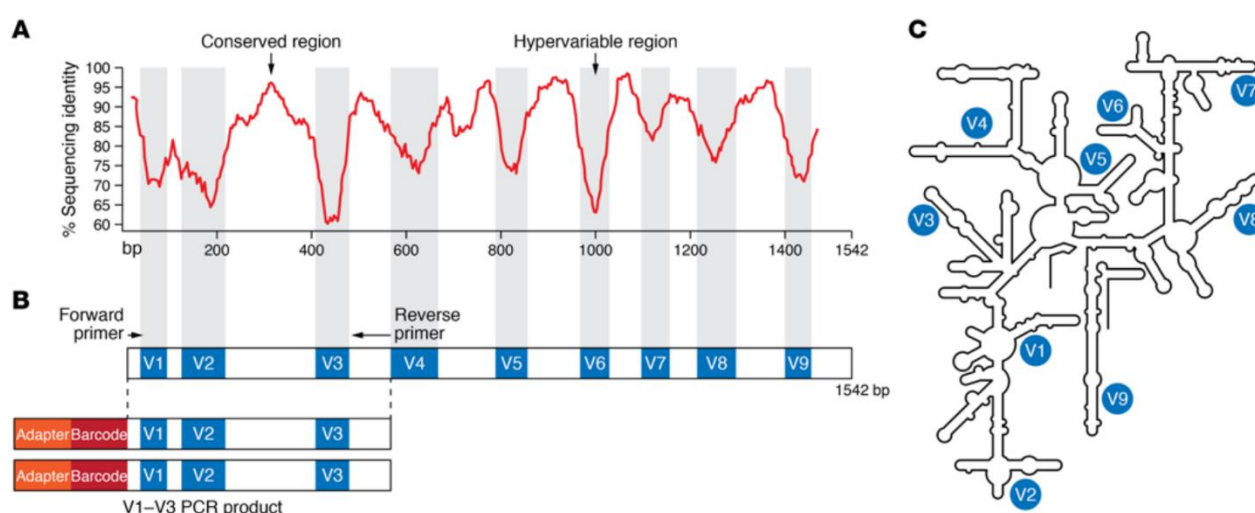


Figure 6. Variable regions of the 16S ribosomal RNA. (A) Percentage sequence identity of conserved and hypervariable regions of the bacterial 16S rRNA gene. (B) Example of PCR amplification of the V1–V3 region of the bacterial 16S rRNA gene. Primers attach to conserved sequences for amplification of hypervariable regions. (C) Schematic structure of the 16S rRNA gene with labelled

hypervariable regions (V1–V9). Adapted from Wensel et al., 2022, permission for figure reproduction granted by the Creative Commons license (<http://creativecommons.org/licenses/by/4.0/>.)

Sequence comparison with annotated bacterial genomes allows for accurate and rapid assignment of the OTUs with their microbial taxonomy (Clarridge, 2004; Pirih & Kunej, 2018). Moreover, analysis of full-length 16S rRNA gene sequencing demonstrates it could be the next step to increase method utility by enabling discrimination of species and strains (Johnson et al., 2019).

Although the microbial world was revolutionized by the analysis of the 16S rRNA gene, new techniques of whole genome sequencing, RNA sequencing, and multi-omics have provided another level of depth in bacterial identification and pushed microbiome studies to higher standards (Figure 7). By sequencing all genetic material from all organisms present in each sample, these new approaches have not only facilitated the study of functional genes contained and expressed by the microbiota, but also led to improved microbe classification at the species level (Galloway-Peña & Hanson, 2020; Wensel et al., 2022).

16S rRNA gene sequencing	Metatranscriptomics
Pros: cost-effective, analytical pipelines widely accepted	Pros: assessment of gene expression
Cons: limitations in species/strains differentiation, lack of functional information	Cons: protein expression may depend on translational modifications
Whole metagenome sequencing	Metabolomics and metaproteomics
Pros: assessment of all bacterial genes within a sample	Pros: assessment of the microbiota functional impact
Cons: computationally demanding, expensive, no consensus on analytical pipelines	Cons: unknown origin of metabolites and proteins

Figure 7. Comparison between different techniques for microbiome analysis. Created with BioRender.com.

For instance, whole metagenome sequencing allows for genome reconstruction, community profiling, and functional profiling of the microbiota through sequencing of random fragments of all DNA in a sample (Figure 8). After DNA extraction, fragmentation and amplification of genetic material generate millions of short DNA sequences, termed reads, which come from all organisms within the community (Figure 8A). Aligning and merging (read assembly) allows reconstruction of longer contiguous sequences (contigs), enabling genome assembly of dominant community members (Figure 8B). Importantly, the number of reads from the genome of one species is proportional to the abundance of that species in the community, but also to the length of its genome, which poses a challenge for detection and assessment of low-abundance species (Joseph & Pe'er, 2021).

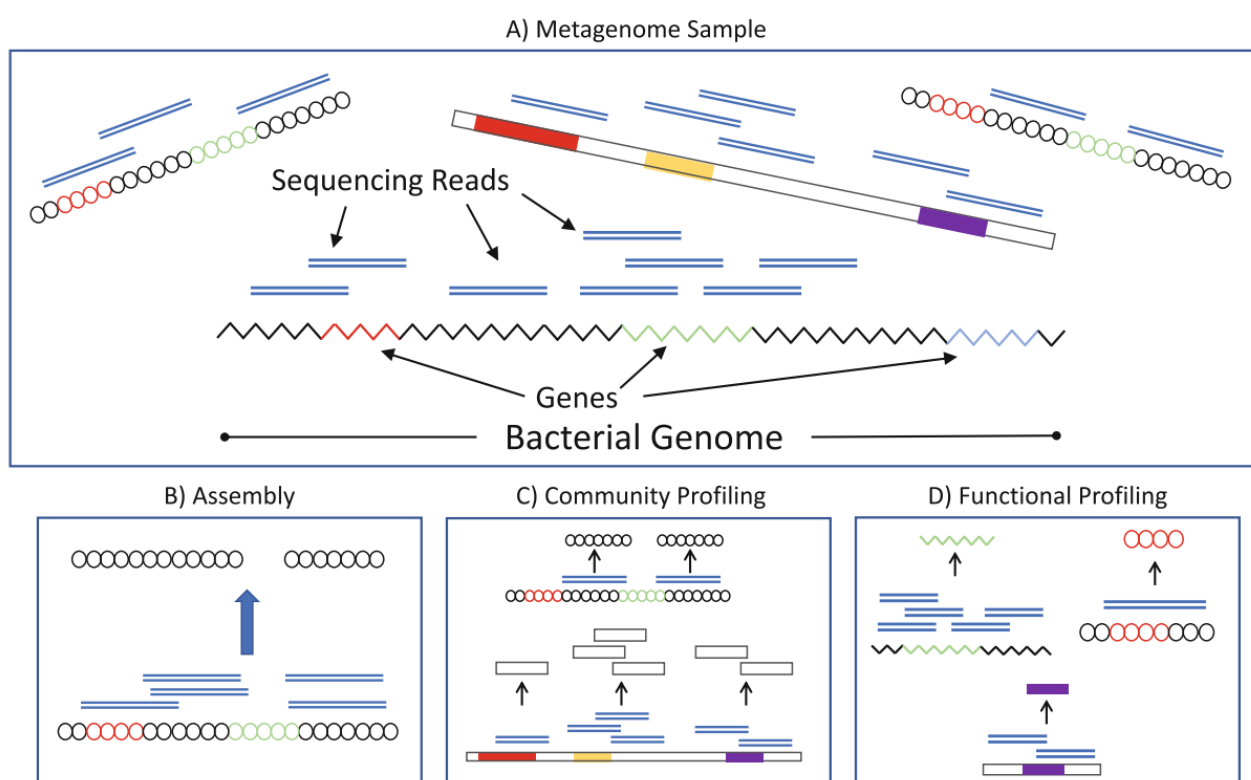


Figure 8. Major components of whole-metagenome sequencing studies. (A) The sampling process for four (linearized) bacterial genomes across three taxa. **(B)** Illustration of genome assembly from sequencing reads. **(C)** Illustration of community profiling, the determination of taxa in a community. **(D)** Illustration of functional profiling, the determination of the functional content in a community. Reproduced from Joseph & Pe'er, 2021 with permission from Springer Nature.

Community profiling (Figure 8C) relies on taxonomic classification tools to identify and quantify the abundance of microbial taxa present in the sample, shedding light on community composition and diversity. Nonetheless, it is limited by the availability of reference genomes, which for many microbes must be reconstructed from metagenomic samples themselves. Thus, novel species lacking a reference genome can only be identified at a higher taxonomic level, such as genus or family, being the accuracy of this classification dependent on how closely related it is to species already present in the database. Additionally, the length of reads significantly influences taxonomic assignment, with shorter read lengths (<200 bp) yielding fewer matches against a reference database and introducing more ambiguity. As shorter reads are more cost-effective, this presents a particular challenge for metagenome studies (Joseph & Pe'er, 2021).

Whereas community profiling focuses on identifying which taxa are contained in a sample, functional profiling (Figure 8D) involves annotating genes within the metagenome to predict their biological functions. Thus, by providing insights into the metabolic potential and ecological roles of microbial populations, functional profiling has the potential to uncover novel molecular targets for drug design by assessing differential gene content between healthy and disease states. Similar to community profiling, read length can also affect the quality and number of annotations in a sample. Additionally, these annotations are also limited by the lack of reference genomes. For instance, the function of uncharacterized genes/proteins may be inferred from homology to previously annotated genes, but robust evidence requires further functional assessment. Lastly, data from less extensively studied systems tend to have fewer functional predictions, and functions shared across multiple taxa are more likely to be annotated (Joseph & Pe'er, 2021).

Importantly, these datasets undergo bioinformatics quality control, including filtering for sequencing errors and removing contaminating sequences through different bioinformatic tools. Subsequently, various bioinformatics tools are employed for taxonomic assignment, functional annotation, and comparative analysis to decipher the intricate relationships and ecological functions of microbial communities within the metagenome. Of note, it has been demonstrated that distinct pipelines may render drastically different results, and the establishment of a consensus approach could help ensure robust interpretations and homogeneity among studies (Nearing et al., 2022).

On the other hand, metatranscriptomics (MTX) refers to the assessment of all transcriptomes (both protein-coding and non-coding RNA) in a given habitat. After establishment of the study group, samples of interest are collected and subjected to isolation of total RNA, which is then inferred by shotgun sequencing as described above. Importantly, microbial and host RNA is unstable if compared to DNA, and cells contained in a sample may respond rapidly to environmental changes. Thus, samples should be preserved and/or processed immediately upon collection to mitigate study limitations. Additionally, as the number of reads from a given mRNA is proportional to its abundance within the sample, it is common that bacterial reads have a decreased reading depth when compared to host reads, and the microbial mRNA portion may be enriched by depleting rRNAs when needed. After sequencing, raw data is converted to functional information and expression estimates using computational analysis. Important steps of this process include subtraction and removal of rRNA reads, taxonomic and functional annotation of reads against reference databases, and read grouping based on their sequence similarity, taxonomic, and/or function information (Figure 9) (Ojala et al., 2023).

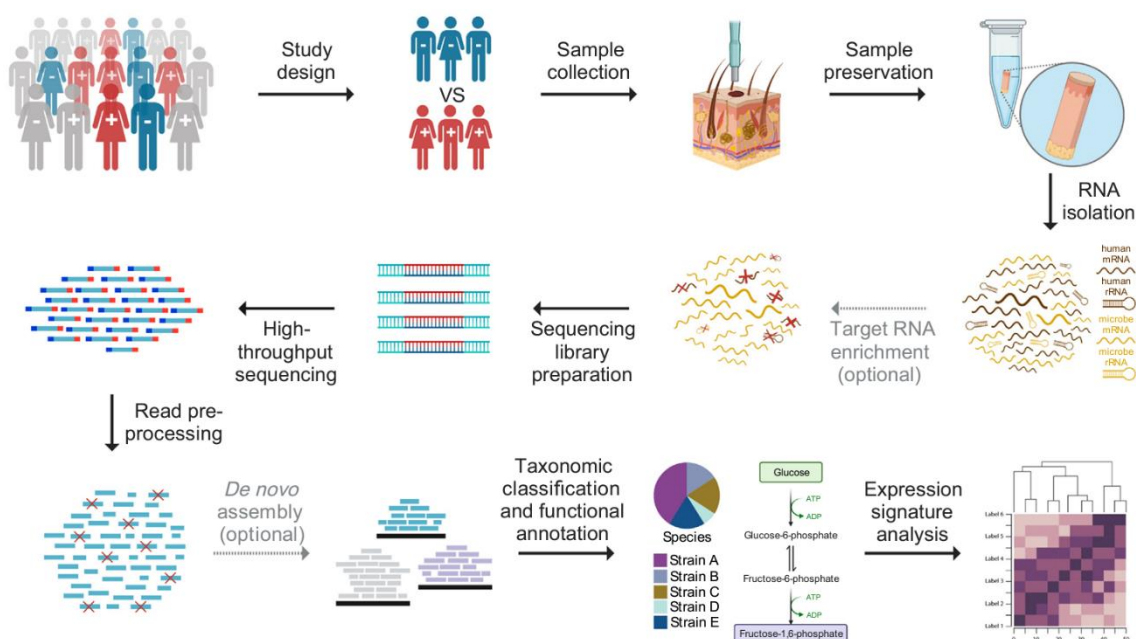


Figure 9. Summary of metatranscriptomics workflow. Data-generation process involves sample collection, RNA extraction, library preparation and sequencing, and bioinformatics processing. Reprinted from Ojala et al., 2023, permission for figure reproduction granted by the Creative Commons license (<http://creativecommons.org/licenses/by/4.0/>.)

While metagenomics provides a static view of the genetic composition of microbial communities, metatranscriptomics assesses the active gene expression patterns, shedding light on the real-time transcribed products within the microbiome and the host. By examining which genes are actively transcribed at a given moment, metatranscriptomics reveals crucial information about the functional repertoire of microbial communities. Moreover, analyses of differential expression between samples enables further understanding of how these communities respond to environmental stimuli and host-microbe interactions. Thus, metatranscriptomics provides one of the most tractable links between the genetic potential of an organism and its molecular activity. This dynamic approach not only enhances our understanding of microbial functionality but also helps identify key pathways and processes that drive microbial community dynamics and impact host health (Ojala et al., 2023; Zhang et al., 2021).

1.2.3 Challenges of studying the endometrial microbiota

The reported lack of consensus on the endometrial microbiota taxonomic core could be due to several limitations in microbiome studies. Unstandardized methodological practices and important differences in study populations make available studies difficult to compare, and invasive sampling methods are major limitations to increasing sample cohort sizes. Taking into consideration the current lack of large cohort studies on the endometrial microbiome, comparability between studies is needed, and the establishment of general protocols would allow this kind of analysis.

1.2.3.1. *Population and Study Design*

The sample cohort is one of the biggest limitations on endometrial microbiome studies. Patient characteristics — such age, sexual habits, and features of the menstrual cycle — should be carefully considered during the process of cohort selection, as several studies have suggested that these characteristics may be responsible for community fluctuations in the uterus and other parts of the FGT (C. Chen et al., 2017; Fujii & Oguchi, 2023; Pelzer et al., 2018; J. Wang et al., 2021). Moreover, important taxa dissimilarities were observed in vaginal samples between patients of different genetic backgrounds, suggesting that Black and Hispanic women harbor a distinct microbiome than Caucasian and Asiatic women (Gajer et al., 2012; Zhou et al., 2007). This may also represent a significant bias in population selection, and further investigation is required to assess this topic on the composition of the endometrial microbiome.

Strict inclusion and exclusion criteria considering hormonal treatment, antibiotic/probiotic administration, and intrauterine devices should also be used because these interventions may influence results through changes in endometrial

microbiome composition (Achilles et al., 2018; Brooks et al., 2017; Carosso et al., 2020; Kadogami et al., 2020). Thus, study designs should consider different demographic and health parameters to mitigate inter- and intraindividual dissimilarity and allow results to be extrapolated to wider populations of infertile patients.

Aside from interindividual dissimilarity, intrapersonal variability may also account for differences among studies. Because there are reports of microbiota changes within the menstrual cycle (C. Chen et al., 2017; Gajer et al., 2012; Kadogami et al., 2020), collection of samples on specific days after menses could also make different studies more easily comparable. For instance, studies assessing the impact of UGT microbiota in ART treatment should acquire samples during the periconceptual period, avoiding bias mainly related to hormonal swings.

1.2.3.2. Sample Handling and Experimental Design

In most studies, endometrial samples are usually collected through the vaginal–cervical route during gynecological assessment or embryo transfers. Thus, there is a risk of contamination with cervicovaginal microbiota during sampling that could impact microbiome analyses (Figure 10), especially considering the low-biomass microbiome found in the uterus and the different bacterial communities found throughout the FGT (C. Chen et al., 2017; Diaz-Martínez et al., 2021; Kyono, Hashimoto, Nagai, et al., 2018; Pelzer et al., 2018).

For instance, in one study only 2 of 26 pairs of samples of endometrial fluid and vaginal aspirates had the same bacterial OTUs, while six paired samples had completely different microbial communities (Moreno et al., 2016). These results suggest that the uterus has its own microbial communities in approximately 20% of

healthy asymptomatic woman, and any kind of contamination from the lower genital tract could result in significant misrepresentation of the endometrial microbiota.

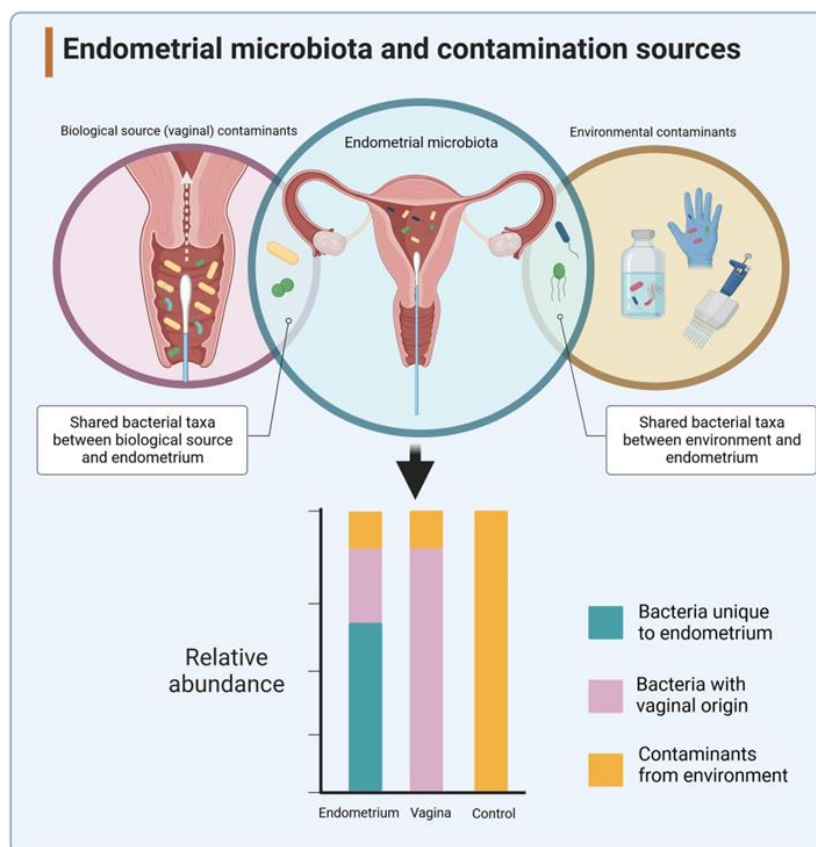


Figure 10. Sources of contamination in endometrial microbiota. Reprinted from Odendaal J. et al., 2024, permission for figure reproduction granted by the Creative Commons license (<http://creativecommons.org/licenses/by/4.0/>.)

Importantly, different microbiome compositions within uterine samples have been observed, and studies suggest that endometrial fluid (EF) and endometrial biopsy (EB) samples might not contain the same bacterial composition. For instance, several taxa could only be found in either EB or EF, including *Achromobacter*, *Brevundimonas*, and Verrucomicrobiaceae (Liu et al., 2018), which could represent actual differences in bacterial composition between these samples or be the results of experimental contamination. Similar results were reported by other groups for minor taxa, as *Streptomyces* and *Clostridium* were detected only in EF, while *Klebsiella* and

Micrococcus were exclusively detected in EB samples. Still, despite different microbiota compositions, dysbiotic states both in EB and EF led to equivalent clinical outcomes (Moreno et al., 2022). These variations in endometrial samples ultimately represent the need for another layer of standardization on experimental design to facilitate comparisons among uterine microbiome studies.

Furthermore, the presence of bacterial DNA in laboratory reagents, plastic consumables, and protective equipment could also significantly impact analyses of bacteria communities in the low-biomass endometrium (Figure 10) (Salter et al., 2014). Although it is difficult to control all sources of contamination, negative and blank controls should be carefully processed to detect basal contaminants, and the use of at least three different negative controls and two positive controls may be important to optimize the detection of contaminants (Molina et al., 2021).

Technical diversity in experimental design can also be considered a source of inconsistency among studies. Different DNA extraction kits, pair of primers used, and sequencing platforms may also render greater differences between studies, and standardized approaches should be considered to facilitate study comparison. For instance, analysis of the V1 region better differentiates *Staphylococcus* species, while V2 and V3 would be more suitable for studying *Mycobacterium* and *Haemophilus* species, respectively (Chakravorty et al., 2007). Importantly, species that impact female health, such as *Gardnerella vaginalis* and *Chlamydia trachomatis*, may not be properly assessed in sequencing based on V1/V2 regions (Graspeuntner et al., 2018). Altogether, these data demonstrate the significant variability in results caused by the choice of different hypervariable regions. As under- or overestimation of taxa impacts analyses of endometrial microbiome diversity, careful considerations should be taken on this matter.

1.3 THE FEMALE REPRODUCTIVE TRACT MICROBIOTA

With the advance of molecular microbiology, the microbiota from the urogenital tract was described to bear around 9% of the total human microbiome (NIH HMP Working Group et al., 2009). Although the uterine cavity was believed to be sterile until the second half of the 20th century when traditional microscopy and culture-based techniques were used to assess the microbiome (Baker et al., 2018; Perez-Muñoz et al., 2017), techniques such as 16S rRNA gene sequencing enabled many studies to describe different microbial communities inside the uterine cavity, yet it is unclear where these originate. For instance, uterine colonization is hypothesized to occur from gut, oral cavity, bloodstream, and vaginal ascension (Baker et al., 2018; Benner et al., 2018), but uterine seeding could potentially occur through assisted reproductive technology (ART) procedures and placement of contraceptive devices (Pereira et al., 2016; Sparks et al., 1981) as well as attachment of microorganisms to human spermatozoa (Svenstrup et al., 2003). Regardless of its origin, research consistently demonstrates that the uterine microbiome is highly diverse and scarcely populated compared with the lower genital tract (Chen et al., 2017; Ichiyama et al., 2021; Wee et al., 2018), but its composition remains to be fully agreed upon.

Significant differences exist between the upper genital tract (UGT) comprising the endometrium and the fallopian tubes; and the lower genital tract (LGT), comprising the cervix and vagina. While the first appears to bear a lower biomass microbiota, containing up to four orders of magnitude fewer bacteria than the LGT, the diversity of bacterial communities in the UGT seems significantly higher (Figure 11) (Chen et al., 2017).

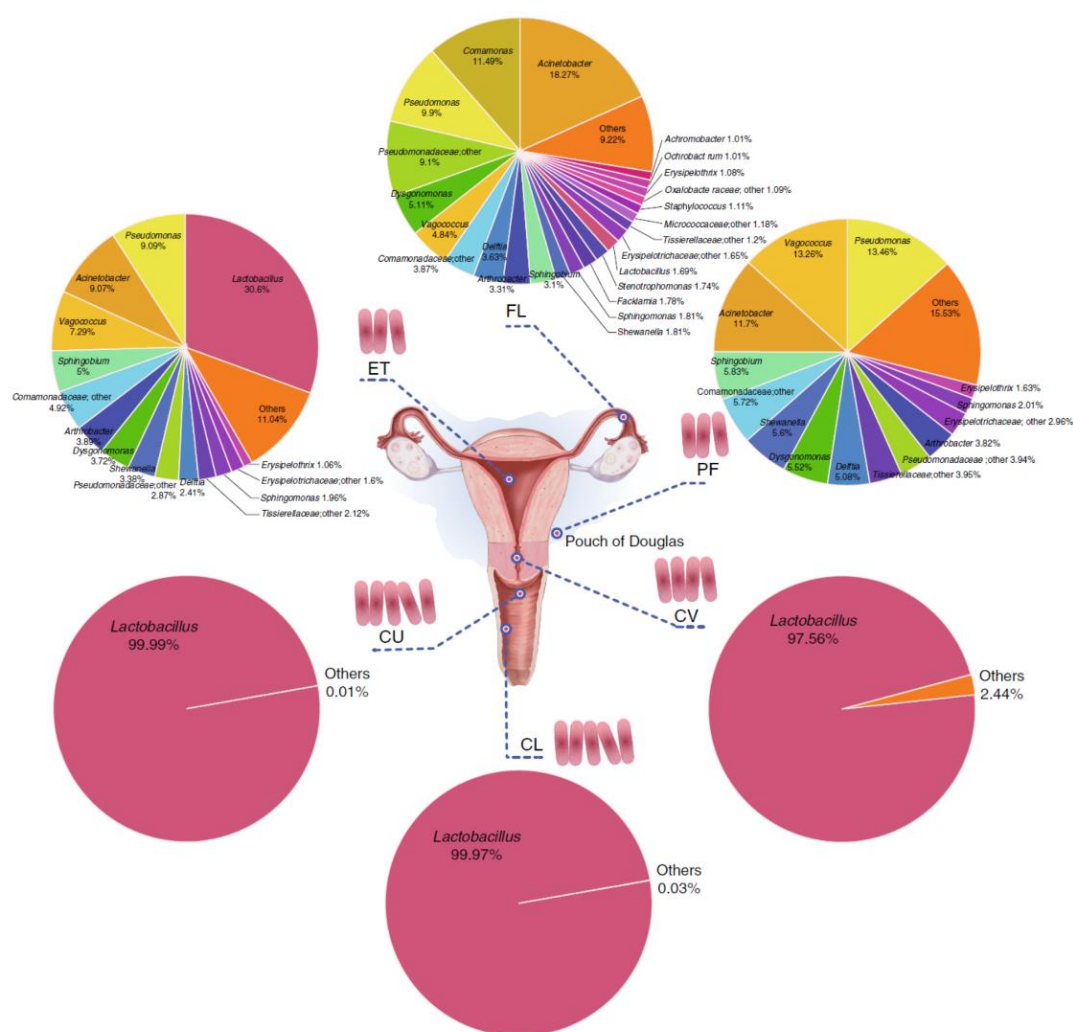


Figure 11. Microbiota composition of the female genital tract. Bacterial taxa are shown according to mean relative abundance. (CL, lower third of vagina; CU, posterior fornix; CV, cervical mucus drawn from the cervical canal; ET, endometrium; FL, fallopian tubes; PF, peritoneal fluid from the pouch of Douglas). Adapted from Chen et al., 2017, permission for figure reproduction granted by the Creative Commons license (<http://creativecommons.org/licenses/by/4.0/>.)

Importantly, metagenomics from the cervical mucus (CV) and the peritoneal fluid (PF) demonstrated not only bacteria, but also fungi, archaea, and virus inhabiting the genital tract of reproductive-age women (Li et al., 2018). Additionally, microbiome functionality revealed peritoneal fluid samples displayed a greater proportion of genes involved in functions such as carbohydrate metabolism, membrane transport, and drug resistance, whereas genes involved in translation, energy metabolism, and metabolism of cofactors and vitamins were enriched in samples taken from the cervical

canal. Moreover, analysis of CV samples revealed an enrichment of transport systems including ABC (ATP-binding cassette) transporters, as well as regulatory systems of aerobic and anaerobic respiration and osmotic stress response (Li et al., 2018). These results demonstrate how different microbiota compositions may have distinct genomic characteristics, significantly affecting community behavior and consequently impacting its interaction with the host.

There are significant fluctuations in bacterial communities in the FGT. These shifts correlate with parameters such as age, hormonal changes, ethnicity, and intrauterine devices (Achilles et al., 2018; Carosso et al., 2020; Chen et al., 2017; Gajer et al., 2012; Wang et al., 2021). Conversely, just like other body sites, alterations in the microbiota composition from the FGT is also implicated in health and disease. In fact, dysbiosis in the FGT has been associated with several gynecological conditions such as BV, chronic endometritis (CE) and endometriosis, but also with embryo implantation, pregnancy establishment and maintenance, and preterm birth (Punzón-Jiménez & Labarta, 2021). Therefore, microbiome assessment and modulation has been proposed as a feasible approach to improve clinical outcomes during assisted reproductive therapy.

The healthy vaginal microbiota has been consistently described as *Lactobacillus*-dominated (LD), as bacteria from this genus represent over 90% of total counts found in vaginal samples from women of reproductive age (Chen et al., 2017). Nonetheless, the vaginal microbiota has been classified into five community state types (CSTs) depending on their bacterial composition (Ravel et al., 2011): while most women presented a vaginal microbiota dominated by *L. crispatus*, *L. gasseri*, *L. iners*, and *L. jensenii* (corresponding to CST-I, – II, – III and V, respectively), women diagnosed with BV generally exhibited a CST-IV vaginal microbiota, with lower percentage of

Lactobacillus spp. and higher levels of anaerobic bacteria such as *Atopobium*, *Dialister*, *Gardnerella*, *Prevotella*, and *Sneathia* (Ravel et al., 2011). Additional assessment revealed the vaginal microbiota could be further divided into thirteen subgroups, depending on the relative abundance of the focal species but also considering minor abundance of other species (Figure 12) (France et al., 2020).

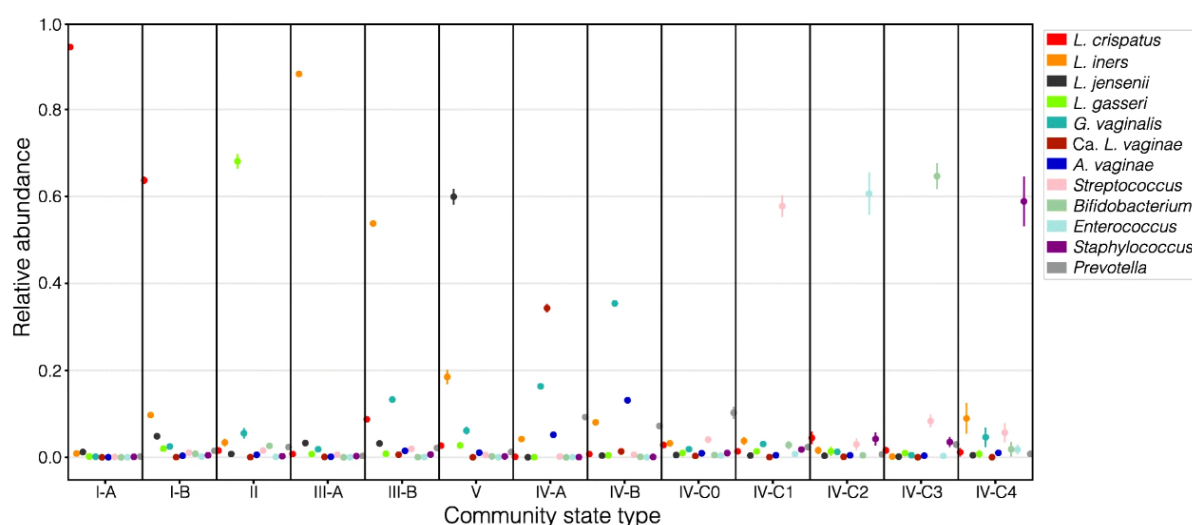


Figure 12. Bacterial composition of the thirteen sub-CSTs described in the vaginal microbiota. Reprinted from France et al., 2020, permission for figure reproduction granted by the Creative Commons license (<http://creativecommons.org/licenses/by/4.0/>.)

The high abundance of *Lactobacilli* has been proposed to play an important role in vaginal health by lowering vaginal pH and producing molecules such as bacteriocins and other substances that inhibit the growth of pathogenic organisms that cause vaginal infections (Sirota et al., 2014), creating a niche that prevents sexually transmitted infections (STIs) and other conditions such as BV (Chee et al., 2020). Moreover, higher abundance of *Lactobacillus* spp. has been consistently associated with increased pregnancy and live birth rates, as well as lower miscarriage rates in women undergoing ART (Souza et al., 2023). Interestingly, in a group of subfertile women at the eight gestational week, vaginal microbiota either remained

Lactobacillus-dominated or transitioned from mixed community to LD, but never from *Lactobacillus* dominance to NLD (Väinämö et al., 2023). These features support the current hypothesis that *Lactobacilli* play an important role in vaginal health and in human conception.

Metagenomics in 82 vaginal samples from healthy women revealed high abundance of *Lactobacillus* spp., including *L. iners*, *L. crispatus*, *L. gasseri*, and *L. jensenii*, as well as common dysbiotic bacteria such as *G. vaginalis*, *Atopobium vaginae*, and *Prevotella amnii*. Moreover, functional annotations demonstrated several pathways were ubiquitous and prevalent across individuals, such as pyruvate fermentation and de novo biosynthesis of adenosine ribonucleotides (Liu et al., 2021). Of note, comparison between women with high risk of preterm birth (n=35) and low-risk controls (n=14) revealed an increased abundance of *L. crispatus* in samples from a full-term pregnancy, whereas *Atopobium vaginae* associated with preterm birth. The preterm birth group also displayed significant associations with genes related to respiratory electron transport chain, alginic acid biosynthetic process, ATP hydrolysis coupled proton transport, and glucose catabolic process, adding further evidence that demonstrate an imbalance in microbe functionality following dysbiosis (Feehily et al., 2020).

1.3.1 Composition of the healthy endometrial microbiota

Traditionally, it was widely agreed that a healthy uterus was sterile, and the presence of bacteria within the uterine cavity was considered pathologic (Baker et al., 2018). However, the hypothesis of the non-sterile uterus and increasing evidence indicating an important role of the uterine microbiota in female health (Toson et al., 2022) suggest that bacterial homeostasis is crucial for IVF success and pregnancy maintenance.

A number of studies describe the healthy state of the uterine microbiota in women of reproductive age, with most reporting dominance of *Lactobacillus* species. Analysis of endometrial fluid samples from 22 fertile patients revealed an abundance of *Lactobacillus* with smaller proportion of other genera such as *Bifidobacterium*, *Gardnerella*, *Prevotella*, and *Streptococcus* (Figure 13) (Moreno et al., 2016). Further assessment prompted a suggestion of two types of bacterial endometrial composition: *Lactobacillus*-dominant (LD; with *Lactobacillus* spp. abundance higher than 90%) or non-LD (NLD; *Lactobacillus* spp. abundance lower than 90%) (Moreno et al., 2016). Moreover, assessment of EF microbiota in seven asymptomatic healthy volunteers revealed that six had high abundance of *Lactobacillus*, which accounted for more than 90% of microbial composition (Kyono, Hashimoto, Nagai, et al., 2018).

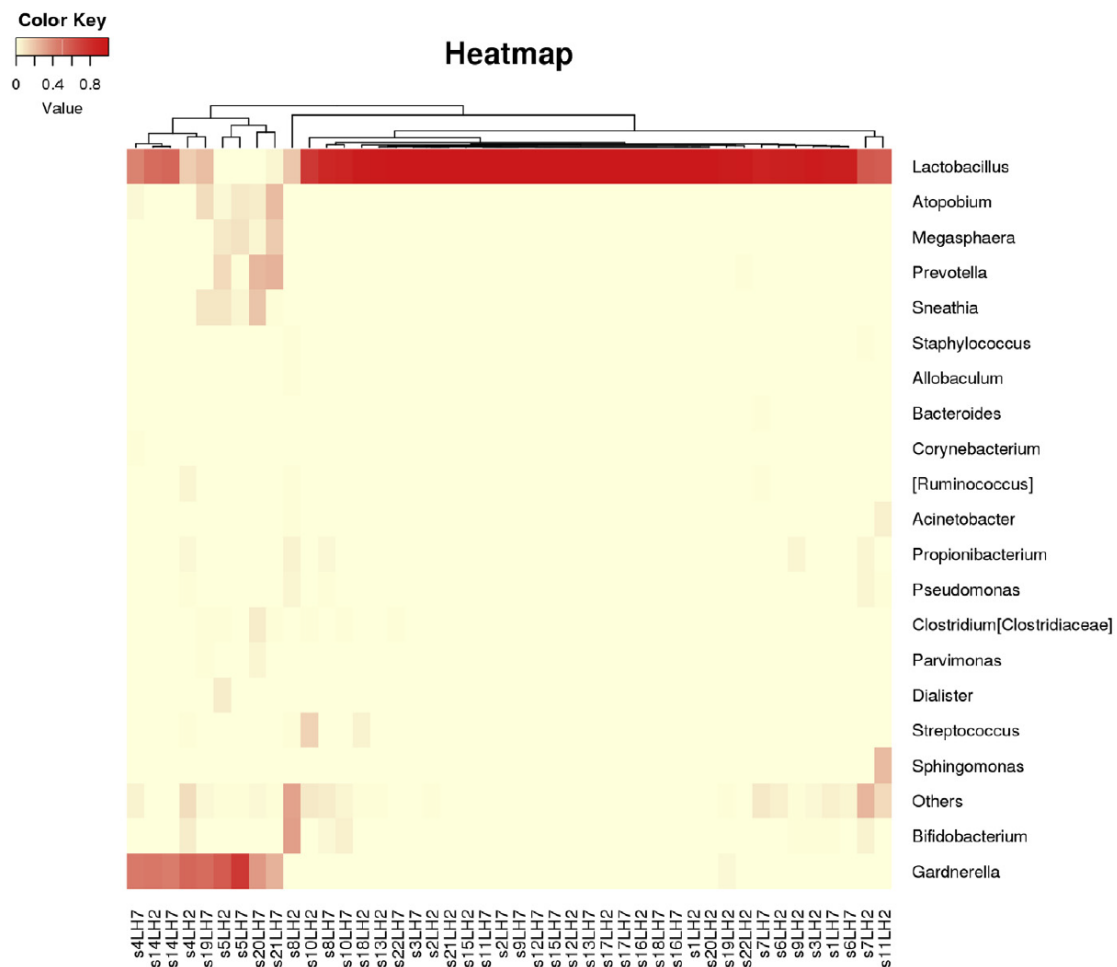


Figure 13. Heatmap depicting endometrial microbiota distribution during the acquisition of endometrial receptivity. A total of 166 different OTUs were identified. The most represented genus was *Lactobacillus* (71.7% of identified bacteria); while *Gardnerella* (12.6%), *Bifidobacterium* (3.7%), *Streptococcus* (3.2%), and *Prevotella* (0.866%) were the other most common genera. Clustering of individual samples showed 2 groups depending on abundance of *Lactobacillus* OTUs: *Lactobacillus* dominated (**LD**)-microbiota and a non-LD (**NLD**). Reprinted from AJOG, 315/6 (684-703), by Moreno et al., Copyright (2016), with permission from Elsevier.

By contrast, some studies did not observe *Lactobacillus* predominance inside the uterine cavity. In 2019, a study was conducted in 19 women of European descent with normal pregnancy at full-term. Analysis of endometrial biopsies revealed a common bacterial composition among subjects, suggesting that genera such as *Acinetobacter*, *Corynebacterium*, *Cutibacterium*, *Escherichia*, *Staphylococcus*, and *Streptococcus* could be part of an endometrial microbiota core. Results also revealed that *Lactobacillus* is present in about 20% of patients and with abundance variability below 16% (Leoni et al., 2019). Moreover, an analysis of uterine microbiota from seven healthy women revealed that endometrial biopsies harbored communities of bacteria (85%), fungi (10%), viruses (5%), and archaea (0.3%); *Clostridium botulinum*, *Hydrogenophaga* sp., *Klebsiella pneumoniae*, and *Pasteurella multocida* were the most abundant microorganisms (Sola-Leyva et al., 2021).

Importantly, several factors have been demonstrated to impact the composition of bacterial communities that inhabit the endometrial cavity. For instance, there is a significant impact of age in bacterial communities inside the uterus. Microbial diversity within samples (alpha diversity) decreases with advancing age, while interindividual similarity (beta diversity) is higher among women ages 20 years and younger, suggesting that dissimilarities among patients accumulate with age (J. Wang et al., 2021).

The endometrial microbiome might also be significantly impacted by hormonal changes. For example, exogenous progestins administration results in loss of

Lactobacillus spp. diversity among women with menorrhagia and dysmenorrhea (n=12); and increased numbers of *Prevotella* spp. and *Sneathia* spp. have been proposed as hallmarks of the proliferative and secretory phases, respectively (Pelzer et al., 2018). Interestingly, the abundance of not only bacteria, but also viruses and archaea, differ significantly between the proliferative and mid-secretory phases in endometrial samples taken in natural cycles (Sola-Leyva et al., 2021). Nonetheless, another study reported the uterine microbiome to be relatively stable during acquisition of endometrial receptivity, as 16S rRNA gene sequencing did not reveal significant differences in bacterial communities comparing endometrial samples collected in the pre-receptive and receptive phases of menstrual cycle (2 and 7 days after the luteinizing hormone (LH) peak) (Moreno et al., 2016).

1.3.2 Implications of endometrial microbiota in human infertility

Several studies have also assessed the composition of endometrial microbiota in different subsets of infertile patients, mainly those suffering from repeated implantation failure, which is often used to describe when two-to-three embryo transfers fail to result in a pregnancy (ESHRE Working Group on Recurrent Implantation Failure et al., 2023). The various causes and the limited understanding of RIF present considerable difficulties in both diagnosing and treating the condition, posing challenges for both patients and healthcare providers (Moustafa & Young, 2020).

In a study including 46 patients enrolled in ART cycles who had a history of RIF (following three or more transfer cycles with good-quality embryos, n=28) or who were attempting IVF treatment for the first time (n=18), those with RIF exhibited a significantly lower microbial diversity in endometrial fluid. Results also demonstrated higher rates of LD microbiota (>90%) and higher numbers of *Gardnerella* in RIF

patients than in the control group, although not statistically significant. Additionally, *Burkholderia* was completely absent from the controls, while it was detected in 25% of RIF subjects (Kitaya et al., 2019). Noteworthy, another study attempted to restore LD endometrial microbiota in patients with a history of RIF (44.9% of which were characterized as NLD, mainly inhabited by *Gardnerella*, *Atopobium*, *Streptococcus*, and *Prevotella*). Interestingly, decreased responsiveness to treatment was reported on those for whom *Gardnerella* was the dominant bacteria (Kadogami et al., 2020).

Later work on the endometrial microbiota in 145 patients with RIF and 21 controls revealed similar alpha and beta diversities between groups. Nonetheless, there were different bacterial compositions, and the RIF group exhibited significant enrichment of 14 genera — including *Atopobium*, *Burkholderia*, *Delftia*, *Gardnerella*, and *Prevotella* — but similar abundance of endometrial *Lactobacillus* spp. (Ichiyama et al., 2021). Greater levels of *L. helveticus*, *Sneathia amnii*, and the genus *Prevotella* were reported among patients with RIF, and higher numbers of *L. iners*, *L. jensenii*, and the genus *Ralstonia* were found in subjects without RIF (Diaz-Martínez et al., 2021). Lastly, a study including 27 women with RIF and 18 controls revealed increased abundance of *Prevotella*, *Dialister*, and *Streptococcus* among patients with successive failed embryo transfers, whereas *Lactobacillus* and *Bifidobacterium* positively associated with the control group (Lozano et al., 2023). Interestingly, dysbiosis marked by decreased numbers of *Lactobacillus* has been related to higher concentrations of inflammatory factors such as IL-6 and IL-1 β in endometrial samples from women suffering from RIF, with additional low levels of anti-inflammatory factors IL-10 and IGF-1. Thus, inflammatory dysregulation triggered by microbiota dysbiosis might have a significant impact in women suffering from this condition (Cela et al., 2022).

1.3.2.1. Endometrial microbiota and reproductive outcomes

Regarding reproductive outcomes, there are controversial results supporting the association of the endometrial microbiota composition with reproductive outcomes after IVF. Some studies performing 16S gene sequencing reported no differences in endometrial microbiota composition between distinct groups. For example, one study reported endometrial microbiome assessments from 33 patients who underwent IVF treatment and transferred a single euploid embryo. NGS of samples collected from the distal part of the catheter demonstrated that *Lactobacillus* and *Flavobacterium* were the predominant bacterial communities in all subjects, regardless of their ongoing pregnancy outcomes (Franasiak et al., 2016). Analysis of endometrial fluid samples from 92 women undergoing IVF treatment demonstrated statistically equivalent pregnancy rates in women with an LD microbiota (considered as $\geq 80\%$ *Lactobacillus*) after antibiotic/probiotic management compared with women with an NDL microbiota ($< 80\%$ *Lactobacillus*), although the chances were slightly higher in the former group (Kyono, Hashimoto, Kikuchi, et al., 2018). Furthermore, an assessment in 99 patients of Asian descent, EF samples were split into groups having eubiotic ($\geq 80\%$ *Lactobacillus* and *Bifidobacterium* spp.) and dysbiotic ($< 80\%$ *Lactobacillus* and *Bifidobacterium* spp. with $\geq 20\%$ of other bacteria) endometrial microbiota. Implantation, miscarriage, and pregnancy rates were similar between groups, suggesting that a dysbiotic uterine environment does not impact fertility (Hashimoto & Kyono, 2019).

Nonetheless, several studies reported contrasting results, and suggested a dysbiotic endometrial microbiota might be associated with reproductive failure. For instance, analysis of EF samples from 35 infertile patients undergoing IVF revealed a decreased abundance of *Lactobacillus* significantly associates with lower

implantation, pregnancy, ongoing pregnancy, and live birth rates (Moreno et al., 2016). Moreover, recruitment of 30 women undergoing ART with frozen embryo transfers revealed that although the *Lactobacillus* genus was not significantly different between reproductive outcomes, *Serratia marcescens*, *Staphylococcus* spp., *Glutamicibacter* spp., and *Delftia* spp. were significantly associated with lower pregnancy rates (Parvanov et al., 2023). Interestingly, a significant change in endometrial microbiota was reported in early pregnancy compared with samples collected before a clinical miscarriage in the same woman. Endometrial fluid samples collected previous to spontaneous miscarriage revealed an NLD profile, in which *Lactobacilli* accounted for 15% of the uterine bacterial composition, and a significant diversity of bacterial genera — such as *Streptococcus*, *Pseudomonas*, and *Propionibacterium* — was present. However, a sample collected at the fourth week of a successful gestation in the same patient showed a significant decrease in bacterial diversity, with *Lactobacillus* dominance accounting for 91% of the sample's microbiota (Moreno et al., 2020).

Adding further evidence about the impact of the endometrial microbiome in reproductive outcomes, a prospective observational multicenter study including 342 infertile patients demonstrated that higher abundance of *Lactobacillus* strongly correlated to LBR, suggesting that *Lactobacillus* could be a biomarker for ART success (Moreno et al., 2022). Additionally, the presence of *Atopobium*, *Bifidobacterium*, *Chryseobacterium*, *Gardnerella*, *Haemophilus*, *Klebsiella*, *Neisseria*, *Staphylococcus*, and *Streptococcus* was associated with worse prognosis after ART treatment. Concretely, *Gardnerella*, *Klebsiella*, and *Streptococcus* were significantly increased in non-pregnant patients, *Klebsiella* and *Staphylococcus* were related to clinical miscarriage, and *Enterococcus* was enriched in subjects that had a biochemical

pregnancy (Figure 14) (Moreno et al., 2022). Grouping samples from all reproductive outcomes, further analysis of co-occurrence bacterial networks showed that *Lactobacillus* was negatively correlated with *Gardnerella*, *Bifidobacterium*, and *Atopobium* and positively associated with the commensals *Clostridium* and *Streptomyces* (Moreno et al., 2022).

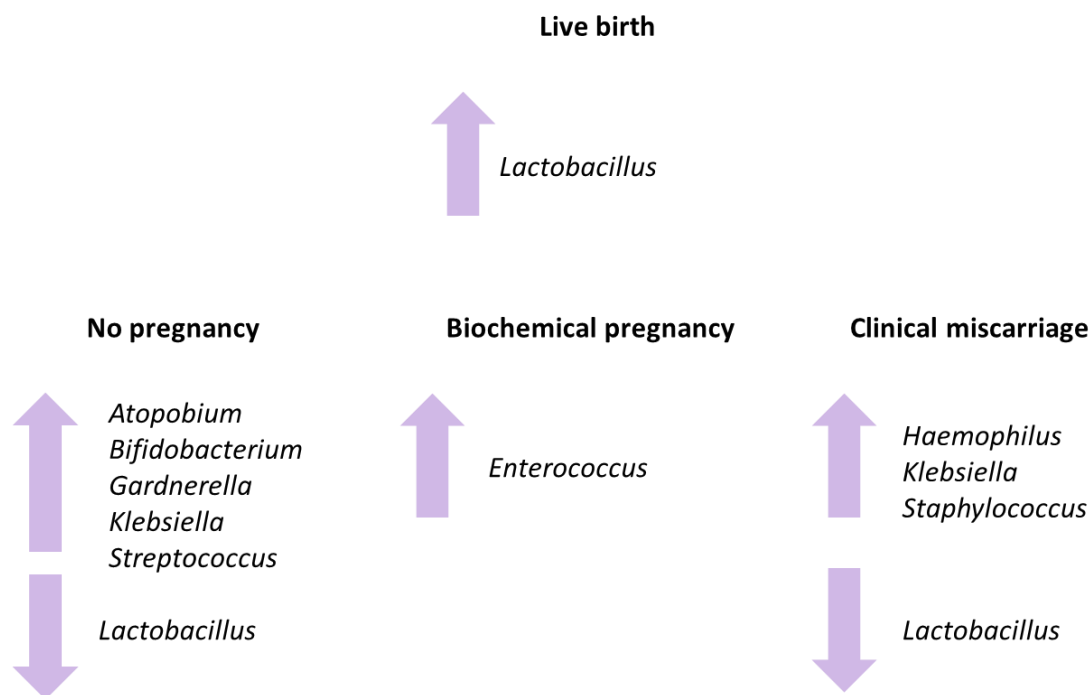


Figure 14. Overview of bacterial genera associated with different reproductive outcomes. Figure created by the author based on results published by Moreno et al., 2022.

Although most of these studies report *Lactobacillus* to be significantly more abundant among patients with better reproductive outcomes, some studies have reported healthy pregnancies in dysbiotic microbiotas. For instance, although a study reported the median percentage of endometrial *Lactobacillus* was significantly decreased in patients undergoing IVF treatment (n=79) compared with healthy volunteers (n=7), seven pregnancies were achieved in NLD cases without any microbiome management intervention, and five of these were ongoing pregnancies by

the time of the report (Kyono, Hashimoto, Nagai, et al., 2018). Similarly, other studies reported cases of normal pregnancies with very low *Lactobacillus* abundance, including patients in which these bacteria were completely absent and other genera such as *Streptococcus*, *Gardnerella*, and *Atopobium* were prevalent (Hashimoto & Kyono, 2019; Leoni et al., 2019; Riganelli et al., 2020).

1.3.3 Implications of endometrial microbiota in gynecological conditions

Increasing evidence demonstrates the impact of variations in the FGT microbiota on overall female health. For instance, several microorganisms related to BV — such as *G.vaginalis*, *A. vaginae*, and *Prevotella bivia* — are associated with the occurrence of CE and pelvic inflammatory disease (Haggerty et al., 2004, 2016; Ravel et al., 2021). Accordingly, endometrial dysbiosis has been associated with the onset of several gynecological conditions, such as CE, endometriosis, and adenomyosis.

Chronic endometritis is characterized by a persistent inflammation of the endometrial lining caused by infection of the uterine cavity mainly by bacterial pathogens such as *Streptococcus* spp, *Escherichia coli*, *Enterococcus faecalis*, and mycoplasma species (*Mycoplasma genitalium* and *Ureaplasma urealyticum*) (Kitaya et al., 2016). Given CE is usually asymptomatic, estimating its real prevalence becomes challenging, with calculations for the general population between 0.8% and 19% depending on the population type and the diagnostic criteria used on each study. Nonetheless, increasing amount of evidence suggests a close relationship between CE and infertility, as its diagnosis has been associated with unexplained recurrent miscarriage or RIF, reaching up to 66% of prevalence in some studies (Kitaya et al., 2016; Moreno et al., 2018). Interestingly, a study including 60 patients undergoing IVF treatment revealed a decrease in *Lactobacillus* spp. abundance among patients

diagnosed with CE through CD138 immunohistochemistry, while *Ralstonia* spp. and *Gardnerella* spp. were associated with CE (Lozano et al., 2021). Lastly, another study included 134 asymptomatic infertile patients receiving ART and reported an endometrial microbiota composed of genera such as *Staphylococcus*, *Gardnerella*, *Atopobium*, and *Fusobacterium* is related to CE (Liang et al., 2023).

Endometriosis is a medical condition in which endometrial tissue grows outside the uterus. The ectopic tissue can be found on the ovaries, fallopian tubes, outer surface of the uterus, and other pelvic organs, causing pelvic pain, abnormal menstrual bleeding, painful intercourse and urination, and even infertility (Giudice & Kao, 2004). Of note, patients suffering from endometriosis have been consistently reported to bear a distinct endometrial microbiota when compared to controls. For instance, among 26 patients with endometriosis and 11 controls, endometriosis patients had a distinct uterine microbiota composition, with enrichment of samples dominated by a mixture of *Acinetobacter*, *Pseudomonas*, and *Vagococcus* in this group (Wei et al., 2020).

Lastly, research reported 64% of patients with endometriosis but <10% of controls had *Fusobacterium* spp. infiltration in the eutopic endometrium. Further assessment revealed that uterine infection by *Fusobacterium nucleatum* significantly induces macrophage infiltration and upregulation of TGF- β signaling in patients with endometriosis. Accordingly, in a mouse model, infection with *F. nucleatum* promotes the growth of endometriotic lesions that can be reverted by specific antibiotic treatment (Figure 15) (Muraoka et al., 2023).

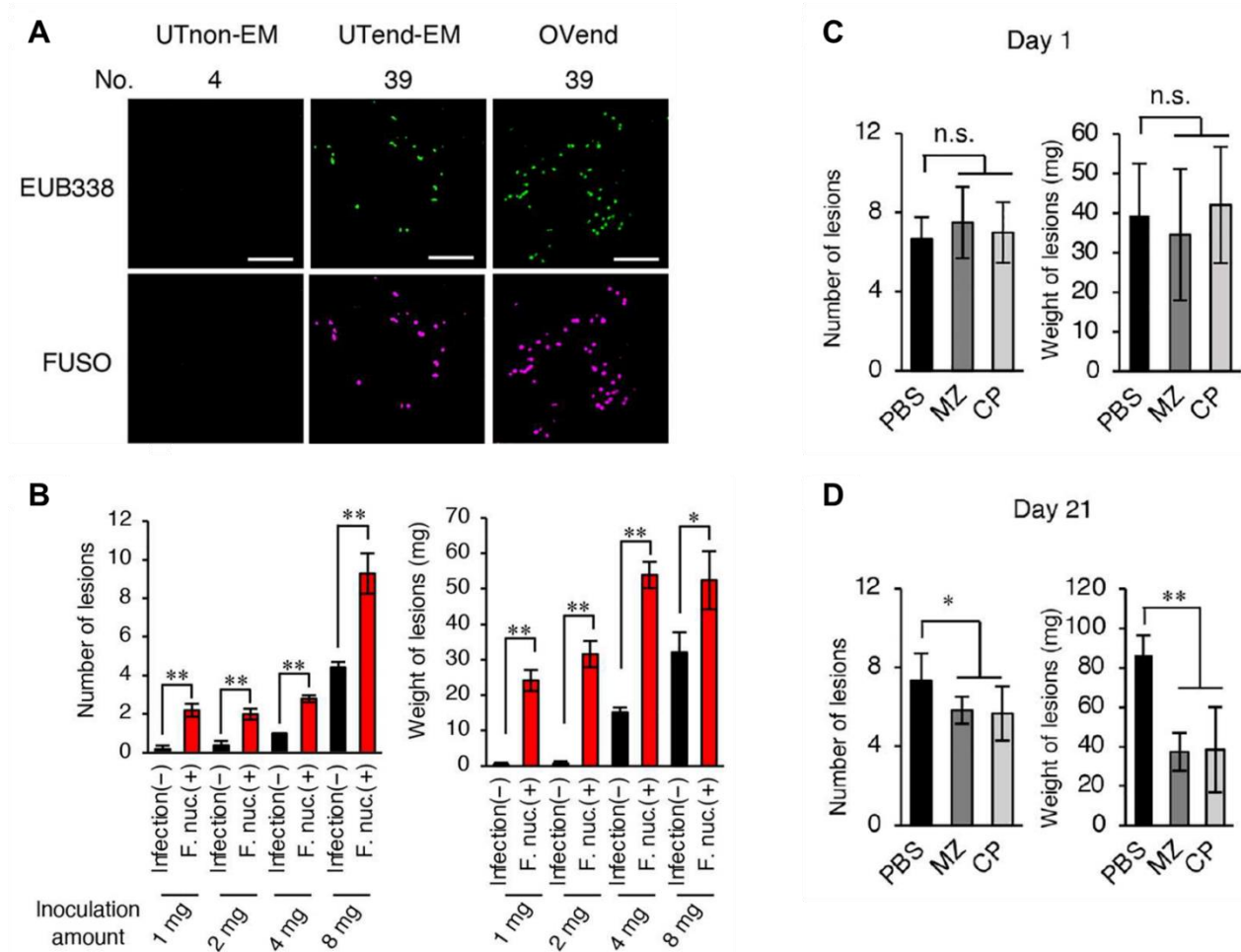


Figure 15. *Fusobacterium nucleatum* and its association with endometriotic lesions in a mouse model. Fluorescence *in situ* hybridization using the EUB338 probe for whole bacteria (green) and FUSO for *Fusobacterium* species (magenta) in uterine and ovarian endometriotic lesions (UTend-EM and OVend) and controls (UTnon-EM) (**A**). Number and weights of endometriotic lesions from recipient mice processed in each group (**B**). Numbers (left) and weights (right) of endometriotic lesions in recipients treated with each antibiotic in each group at day 1 (**C**) and day 21 (**D**). Abbreviations: **MZ**, metronidazole; **CP**, chloramphenicol. Adapted from Muraoka et al., 2023 with permission from the publisher.

Similar to endometriosis, adenomyosis is a medical condition that occurs when the endometrium starts to grow into the muscular wall of the uterus (myometrium). This can lead to enlargement of the uterus and can cause various symptoms such as uterine bleeding, pelvic pain, heavy menstrual flow, and infertility (Bourdon, Santulli, Marcellin, et al., 2021). A study profiling eutopic endometrial samples from 38 women with (n=21) or without (n=17) adenomyosis also reported some possible biomarkers,

describing the presence of *Citrobacter freundii*, *Prevotella copri*, and *Burkholderia cepacia* as potential biomarkers for adenomyosis: at the genus level, the diagnostic sensitivity and specificity for adenomyosis were 71.4 and 100.0%; while the diagnostic sensitivity increased to 81% at the species level (Q. Lin et al., 2023). Moreover, functional profiling revealed significant differences between groups, as controls exhibited upregulation in pathways such as environmental information processing, secretion system, ABC transporters, and other ion-coupled transporters, while patients with adenomyosis displayed upregulation in genes associated with nucleotide metabolism, replication and repair, translation, and others (Lin et al., 2023).

Considering the impact of such conditions on fertility, further understanding of how swings in endometrial microbiota are associated with these pathologies may also provide important insights into new ART treatments through microbiota management.

1.3.4 Tools for endometrial dysbiosis management

Among methods to modify the FGT microbiota, antibiotics have the greatest number of available reports. For instance, antibiotics such as secnidazole and amoxicillin were reported to have a great impact on gynecological pathologies, as treatment led to reduced risk of BV, CE resolution, and higher pregnancy and live birth rates (Abd El Aziz et al., 2019; Cicinelli et al., 2014, 2015, 2018; Iwami et al., 2023; Zou et al., 2023), and animal studies suggest that antibiotic treatment could be a potential novel approach in combination with conventional protocols to treat endometriosis (Kitaya & Yasuo, 2023; Muraoka et al., 2023)..

Even though treatment of these conditions could improve ART outcomes, administration of antibiotics alone as a first choice to alter the microbiota remains controversial, since broad-spectrum antibiotics could impact not only pathogenic

bacterial populations but also the proliferation of protective microbiota such as *Lactobacillus*. Additionally, the increasing challenge of antibiotic resistance may impair treatment efficacy and lead to poor reproductive outcomes. A meta-analysis evaluated the effects of antibiotics on clinical outcomes among patients with CE and showed that patients with cured CE had improved live birth, ongoing pregnancy, clinical pregnancy, and implantation rates compared with patients with persistent CE. Additionally, IVF results were comparable between women with cured CE and no-CE controls (Vitagliano et al., 2018). However, in the same study, the authors did not observe any significant improvement in patients without confirmation of a CE cure, suggesting that CE resolution should always be confirmed before proceeding with ART.

Lastly, Iwami and colleagues conducted the first interventional study, which offered endometrial microbiota assessment (and personalized antibiotic treatment, if needed) to RIF patients as an add-on to their infertility treatment and then compared reproductive outcomes with patients who refused endometrial microbiota testing (Iwami et al., 2023). This prospective study revealed a significant increase in clinical pregnancy, ongoing pregnancy, and live birth rates after the first frozen embryo transfer in patients who agreed to the assessment and personalized treatment based on the endometrial microbial composition analyzed using 16S rRNA gene sequencing in the case of pathogenic endometrial microbiota. These findings suggest that an optimal endometrial microbiota remains vital to improving therapeutic strategies in RIF patients (Table 1).

	Study group (tested by EMMA)	Control group (not tested by EMMA)	p value
Transferred embryos (n)	1.36 ± 0.54	1.34 ± 0.48	0.844
Endometrial preparation for embryo transfer			
Hormone cycle	127 (96.9%)	63 (98.4%)	1.000
Natural cycle	4 (3.1%)	1 (1.6%)	
Results after the first FET following the proposed interventions			
hCG positive rate	62.6% (82/131)	14.1% (9/64)	< 0.001*
Clinical pregnancy rate	42.0% (55/131)	12.5% (8/64)	< 0.001*
Ongoing pregnancy rate	33.6% (44/131)	9.4% (6/64)	< 0.001*
Early miscarriage rate	20.0% (11/55)	25.0% (2/8)	0.665
Live birth rate	33.6% (44/131)	9.4% (6/64)	< 0.001*
Results after 2 cumulative FET cycles			
Implantation rate	48.3% (87/180)	33.7% (29/86)	0.025*
Clinical pregnancy rate	64.5% (79/131)	33.3% (25/64)	0.005*
Ongoing pregnant rate	48.9% (64/131)	32.8% (21/64)	0.028*
Multiple pregnancy rate	9.4% (6/64)	5.0% (1/20)	0.680
Biochemical pregnancy rate	21.4% (28/131)	10.9% (7/64)	0.111
Early miscarriage rate	13.0% (15/79)	11.8% (4/25)	0.498
Stillbirth rate	0% (0/64)	4.8% (1/21)	0.247
Live birth rate	48.9% (64/131)	31.2% (20/64)	0.020*
Weeks of gestation for single births	38.8 ± 1.71	37.6 ± 3.42	0.044*
Single birth weight (g)	2995.8 ± 380.7	2816.8 ± 688.5	0.157

Table 1. Pregnancy outcomes after frozen embryo transfer in patients with or without personalized treatment recommendations based on the microbial 16S rRNA gene sequencing. Data are presented as the mean ± SD unless otherwise specified. Abbreviations: EMMA endometrial microbiome metagenomic analysis, **FET** frozen embryo transfer, **hCG** human chorionic gonadotropin. * p < 0.05. Reproduced from Iwami et al., 2023 with permission from Springer Nature.

Administration of probiotics to induce recolonization of the genital tract has been discussed as a complementary treatment approach. For instance, in patients with BV, probiotic supplementation with *Lactobacillus* led to better clinical outcomes compared to treatment with metronidazole alone (Recine et al., 2016). Additionally, the administration of oral and vaginal probiotics successfully increased the *Lactobacillus* content in endometrial samples from patients with RIF; and patients treated with a combination of antibiotics and vaginal probiotic suppositories had the highest cure rates amongst other treatment strategies (Kadogami et al., 2020).

Despite challenges associated with uterine recolonization through oral administration of probiotics, mainly due to the low pH of the gastrointestinal system (López-Moreno & Aguilera, 2021), treatment with vaginal probiotic suppository in combination with antibiotics led to more significant results if compared to the oral administration route, and this approach could significantly improve clinical outcomes in patients with RIF (Kadogami et al., 2020). Although strong evidence supports an increased abundance of *Lactobacillus* species can be achieved after probiotic administration, there is a lack of homogeneity in treatment protocols, as time of treatment may vary from one day to several weeks, and doses may be significantly distinct (López-Moreno & Aguilera, 2021). Thus, systematic reviews have reported controversial results regarding the effectiveness of oral and vaginal probiotics in improving outcomes for women undergoing fertility treatment (Corbett et al., 2021; Jarde et al., 2018), and standardization of trials would benefit study comparison and facilitate the optimization of this treatment approach.

Promising results with fecal microbiota transplantation (FMT) indicate that this approach may also carry significant therapeutic potential for various conditions ranging from gastrointestinal and autoimmune diseases, cancer, brain disorders, and metabolic syndromes (Biazzo & Deidda, 2022). These results prompted the use of vaginal microbiota transplantation (VMT) as an emerging intervention to treat imbalanced vaginal microbiota by restoring its overall diversity (Tuniyazi & Zhang, 2023). In fact, VMT has been successfully used to treat bacterial vaginosis, with 80% of full long-term remission after treatment (Lev-Sagie et al., 2019). Importantly, VMT in a rat model (containing *Lactobacillus helveticus*, *Lactobacillus crispatus*, *Lactobacillus acidophilus*, *Lactobacillus gasseri*, and *Lactobacillus salivarius*) was demonstrated to not only restore vaginal diversity but also reduce bacterial-derived

inflammation in the uterine walls (Chen et al., 2021). Ultimately, it could be speculated that similar results would apply to other endometrial conditions, and studies should be conducted to determine key parameters — such as dosage, method of application, and best microbiota composition — to select healthy donors and validate its clinical application.

Although further research regarding the clinical applications of VMT is still needed, gathered results suggest it could be an effective treatment for gynecological and obstetrical pathologies in patients who are unresponsive to standard antibiotic/probiotic protocols. Notably, the absence of published studies on endometrial microbiota transplantation can be attributed to challenges associated with its low biomass, besides the invasive nature of sample collection that poses a challenge for donor recruitment. Nonetheless, other strategies such as diet, phage therapy, and postbiotics may represent promising management approaches (Hitch et al., 2022), and its role in the treatment of gynecological conditions should be further assessed.

1.4 GENETIC AND FUNCTIONAL DIVERSITY WITHIN THE MICROBIOME

With increasing advances in molecular biology and sequencing techniques, examining entire bacterial genomes represents a significant improvement in gaining a comprehensive understanding of microbial dynamics, surpassing the limitations of solely analyzing the composition and abundance of bacterial communities. In fact, traditional approaches often focus on identifying species abundance, overlooking the nuanced genetic variations within strains, while new sequencing technologies enable detailed assessment of complete bacterial genomes and unveil essential information about genetic traits, including virulence factors and antibiotic resistance genes. Thus,

metagenome studies will enable further understanding the endometrial microbiome and its functional impact in human reproduction.

1.4.1 Antibiotic resistance

The accidental discovery of penicillin in 1928 by Alexander Fleming revolutionized medicine, marking the advent of antibiotics that transformed the landscape of infectious disease treatment (Durand et al., 2019). However, over time, the misuse and overuse of antibiotics have accelerated the development of resistant strains of bacteria, increasingly affecting the efficiency of these drugs. The rise of antibiotic-resistant bacteria poses a serious threat to public health, as common infections become harder to treat, leading to prolonged illnesses, higher healthcare costs, and an increased risk of mortality; but also modifying the equilibrium and composition of environmental microbial communities (Baquero, 2021).

The resistome comprises the collection of antibiotic resistance genes (ARG) that confers antimicrobial resistance, and its diversity is influenced by various environmental factors including diet and antibiotic exposure (Singh et al., 2019). Importantly, bacteria can acquire resistance genes through horizontal gene transfer, which enables the rapid transmission of resistance traits between different bacterial strains. This share of genetic material with neighboring bacteria can occur through mechanisms like plasmids or transposons, fostering the rapid dissemination of resistance (Lerminiaux & Cameron, 2019). The resistome includes ARGs that act via various mechanisms employed by bacteria to diminish drugs' efficacy. These mechanisms may be separated into two big groups: those that act directly on the antibiotic molecule, and those whose action interferes with the drug target site (Munita & Arias, 2016) (Figure 16).

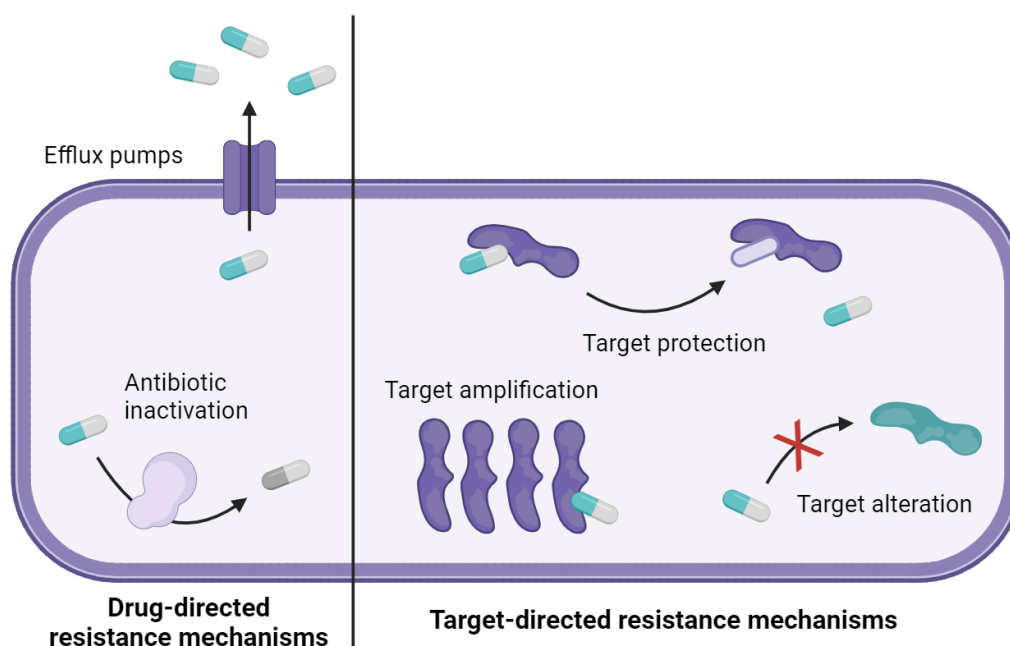


Figure 16. Examples of antibiotic resistance mechanisms. Created with BioRender.com.

Among the drug-directed resistance mechanisms, decreased cell permeability and the ability to pump out antibiotics are important sources of antibiotic inefficiency by preventing sufficient drug concentrations from reaching a therapeutic effect. Additionally, modifications of the antibiotic molecule by acetylation, phosphorylation, or adenylation may change its chemical structure and reduce its affinity for the target site. Lastly, the production of β -lactamases is another widely common resistance mechanism, which destroys the amide bond of the β -lactam ring and inactivates antimicrobial function (Munita & Arias, 2016). Target-directed mechanisms, on the other hand, involve modification, amplification, or protection of antibiotics target sites. These alterations in the target site may impair proper binding between the drug and its receptor, but also result in decreased affinity for the antibiotic molecule, ultimately weakening treatment efficacy (Munita & Arias, 2016).

The use of antibiotics plays a crucial role in different aspects of reproductive medicine, serving to prevent and treat infections that may impact fertility and reproductive health (Cheng et al., 2022; Dalby & Stoner, 2022; Tomás et al., 2020). Primarily, antibiotics are commonly prescribed as a prophylactic measure to reduce the risk of infection during procedures like oocyte retrieval, surgical sperm retrieval, and embryo transfer during ART (Khoa et al., 2022; Sallam et al., 2022). However, additional findings support that therapeutic interventions based on gene sequencing analysis of microbial 16S rRNA in the intrauterine microbiota significantly improves pregnancy outcomes in women undergoing ART (Iwami et al., 2023). Thus, while antibiotics are valuable tools in reproductive medicine, their conscious use is essential to prevent the development of antibiotic resistance, which could significantly impact overall reproductive health and clinical outcomes of patients. Additionally, it is important to consider that some antibiotics may indiscriminately kill beneficial bacteria, potentially disrupting the delicate balance of the microbiota and affecting reproductive health.

The interconnected nature of our world facilitates the rapid spread of resistant strains globally, which highlights the growing threat of antibiotic resistance that could also affect reproductive health and human fertility. Thus, innovative research for new antimicrobial agents and a collective commitment to responsible antibiotic use across healthcare and everyday life are critical to safeguard antibiotic effectiveness and overall patient health.

1.4.2 Virulence factors

Virulence factors are specific characteristics possessed by a pathogen that enhance its ability to establish a successful colonization in a host organism, possibly

causing a disease in a susceptible host. Importantly, they have a wide range of attributes, including the ability to invade the host (exoenzymes), to evade and counteract host defenses (immune modulators), to acquire iron and nutrients from the environment (transporters and nutritional factors), to sense environmental changes (two-component system and quorum sensing), to adhere to surfaces and induce biofilm formation (fimbriae), among others. Taken together, the presence of virulence factors expressed by microbes is tightly associated with the pathogen's ability to establish and maintain infection (Casadevall & Pirofski, 2009) (Figure 17).

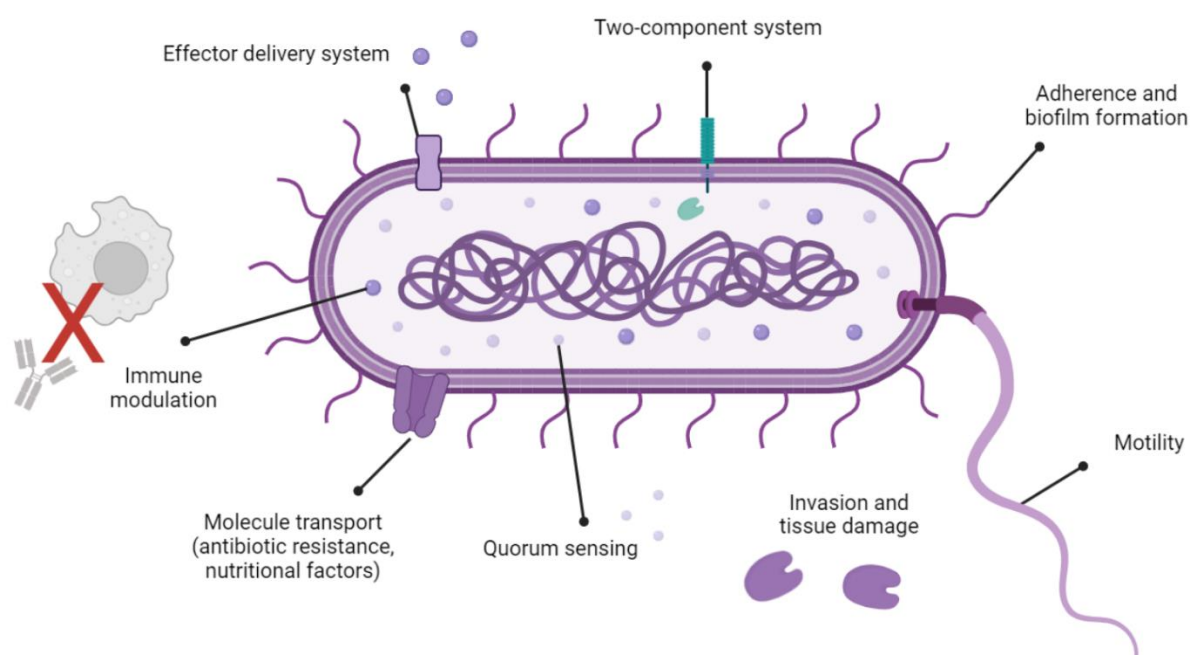


Figure 17. Examples of virulence factors in a bacterial cell. Created with BioRender.com.

The two-component systems (TCS) and quorum sensing (QS) are two types of virulence factors mainly related to environmental adaptation. The two-component systems act at individual cell level and include processes that allow bacteria to adapt to environmental stimuli such as changes in nutrient availability, temperature, and pH. It generally comprises a pair of proteins, namely the sensor kinase and the response

regulator, that sense external stimuli and alter the expression profile of bacterial genes for survival and adaptation, respectively (Hirakawa et al., 2020). On the other hand, quorum sensing may be defined as the cell-to-cell communication mechanism by which bacteria can sense their population density by the production of small molecules. It acts on three distinct phases, from production of small signaling, to its accumulation depending on bacterial density, and to a specific bacterial response when a threshold concentration is reached (Casadevall & Pirofski, 2009). Thus, QS involves population-wide coordination and is commonly associated with the regulation of behaviors such as biofilm formation and secretion of bacteriocins to promote competition against foreign species (Abisado et al., 2018).

Biofilm formation is one of the most clinically relevant among several virulence factors. By producing an extracellular matrix composed mainly of polysaccharides, proteins, and nucleic acids, single planktonic bacteria adhere to surfaces and encase themselves, creating complex and organized communities that characterize biofilms (Rather et al., 2021) (Figure 18).

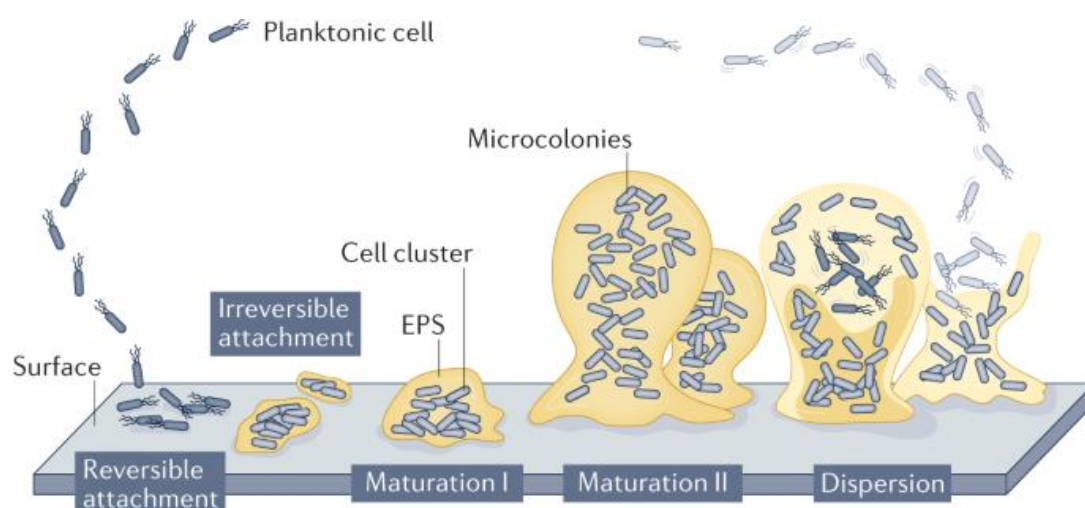


Figure 18. Model of biofilm formation. Abbreviations: **EPS**, extracellular polysaccharides. Reprinted from (Sauer et al., 2022) with permission from Springer Nature.

This matrix provides structural support and protection to the microbial community, allowing it to adhere firmly to diverse surfaces. Nonetheless, these features are an increasing problem for public health, as biofilm formation on abiotic surfaces in food and medical sectors may lead to foodborne, wound chronicity, recurrent and chronic infections, and general antibiotic resistance (Abdallah et al., 2014; Goswami et al., 2023; Sharma et al., 2019). The protective nature of the biofilm matrix makes these communities challenging to eradicate, highlighting the importance of understanding biofilm formation and function for developing effective strategies in medical, environmental, and industrial settings.

Bacteria have also evolved a range of sophisticated strategies to modulate and evade the host immune system, ensuring their survival and proliferation within the host. One common mechanism is through the alteration of surface molecules, such as lipopolysaccharides (LPS) in Gram-negative bacteria, which can be modified to reduce recognition by the host's immune system. By altering these surface components, bacteria may avoid detection by pattern recognition receptors (PRRs) like Toll-like receptors (TLRs), thereby preventing the activation of innate immune responses (Matsuura, 2013). Additionally, some bacteria produce proteases that degrade host immune signaling molecules, such as cytokines and chemokines, effectively dampening the immune response and allowing the bacteria to persist in the host environment (Potempa & Pike, 2009). Furthermore, many pathogenic bacteria produce a polysaccharide capsule that encases the cell, providing a physical barrier that inhibits phagocytosis by immune cells and prevents complement activation, thus protecting the bacteria from immune clearance (Hsieh & Allen, 2020). Lastly, certain bacteria can invade and survive within immune cells like macrophages, effectively hiding from extracellular immune attacks. For instance, *Mycobacterium tuberculosis*

can inhibit the fusion of phagosomes with lysosomes within macrophages, preventing their destruction (Zhai et al., 2019). These immune evasion tactics highlight the dynamic interplay between pathogenic bacteria and the host immune system, illustrating the bacteria's ability to manipulate host defenses to establish and maintain infections.

Another relevant example of bacterial factor that promotes virulence is the secretion of proteins and other molecules from the bacterial cell to its external environment or directly into host cells. In fact, secretion systems have been implicated in virulence in various ways, from enhancing attachment to eukaryotic cells, to gathering resources in an environmental niche or directly intoxicating target cells and disrupting their functions (Green & Mecsas, 2016). Importantly, bacterial secretion systems are diverse and can be categorized based on their unique structures, functions, and substrate specificities. While some secretion apparatuses are universally present across various bacterial classes, capable of transporting a wide range of substrates, others are more specialized, existing only in a limited number of bacterial species and often tailored to secrete a specific set of proteins. This variability allows bacteria to adapt their secretion mechanisms to their specific ecological niches and pathogenic strategies (Coulthurst, 2019; Green & Mecsas, 2016; Schindler et al., 2023; Wagner et al., 2018). Thus, secretion systems are also important for microbe competition for resources and space and may have a significant impact on shaping community composition and structure.

Understanding the role of virulence factors is crucial in the study of infectious diseases, as it helps researchers and clinicians develop novel treatment strategies. Considering the current problem of drug resistance/tolerance, developing innovative antimicrobial therapeutic approaches that could reduce the burden of bacterial

virulence without causing cell death may be of great interest. In fact, next-generation antimicrobials (NGAs) represent a hallmark in the ongoing battle against infectious diseases. These compounds have antivirulence properties at concentrations that do not impact bacterial viability, disrupting the pathogenic potential of bacteria and impairing their ability to establish a colonization, which ultimately prevents disease from happening (Gadar & McCarthy, 2023). Additionally, by disabling key mechanisms of infection establishment and maintenance, these alterations could make microbes more vulnerable to clearance by the immune system and potentially render them more susceptible to traditional antibiotics (Gadar & McCarthy, 2023).

NGAs might target different intra or extracellular components, acting on several virulence mechanisms. For example, the use of distinct NGAs may inhibit adhesion to the host cell, lead to reduced expression of biofilm-related genes, and disrupt quorum sensing and two-component system pathways (Figure 19). The efforts on NGA development aim to overcome the limitations of existing treatments, enhance efficacy, and minimize the emergence of resistance. Nonetheless, the capacity of bacteria to overcome their action still needs to be further explored, and the possible effects on commensal bacteria from the healthy microbiome should also be taken into consideration. Still, targeting virulence rather than microbe viability has immense potential in human and animal health, and the use of NGAs in vaccines or antimicrobial therapies is likely to have increased clinical importance in the future.

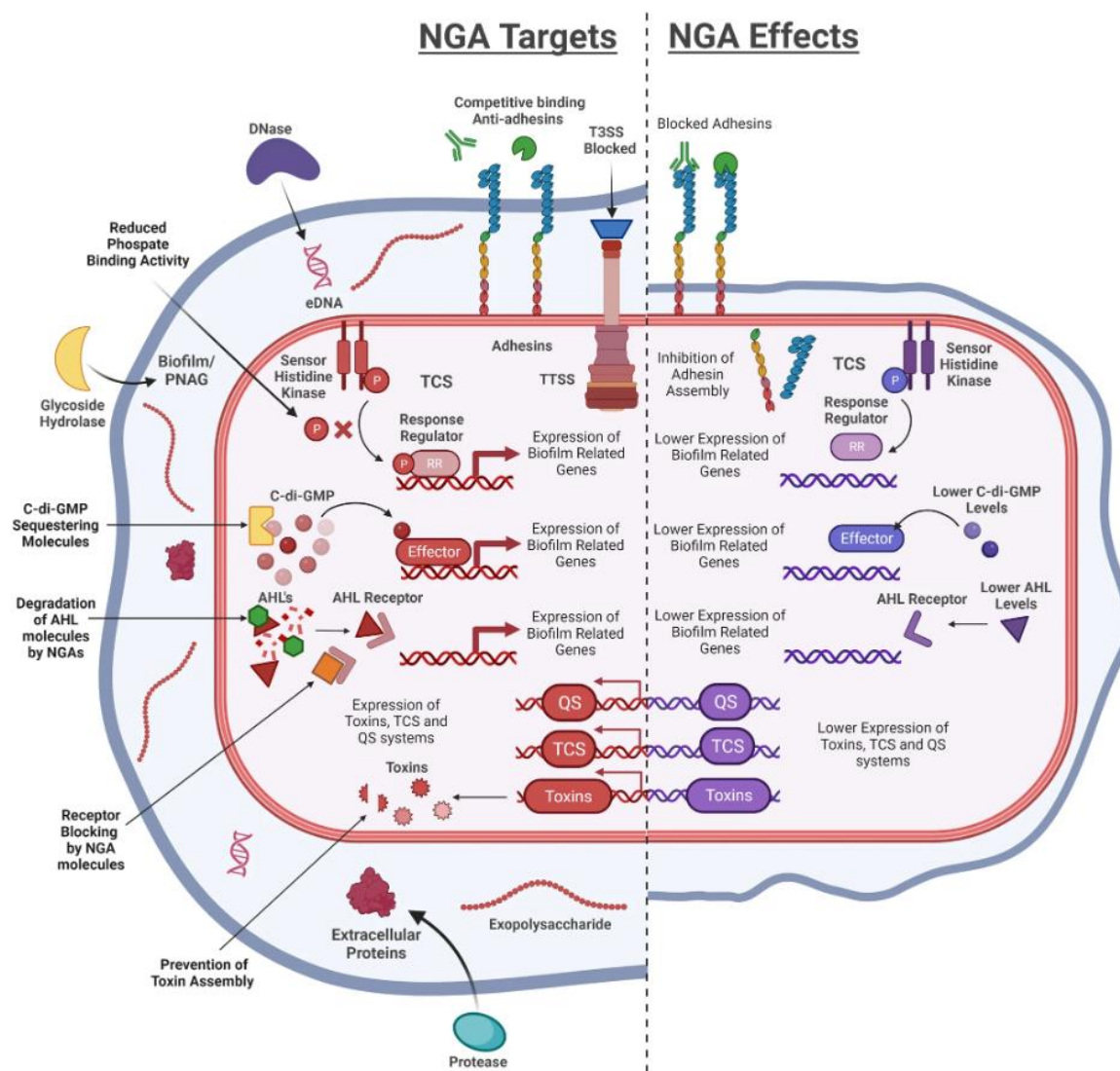


Figure 19. Different cellular and extracellular targets of next generation antimicrobials (NGAs) and the impact on different virulent phenotypes. Abbreviations: **QS**, quorum sensing; **TCS**, two-component system; **T3SS**, Type-III secretion system. Reprinted from Gadar K and McCarthy R, 2023, permission for figure reproduction granted by the Creative Commons license (<http://creativecommons.org/licenses/by/4.0/>.)

1.4.3 Microbial factors and host cell modulation

Bacterial-derived metabolites have been consistently reported to modulate human tissue homeostasis, mainly in the gut. For instance, microbiome-derived stimuli such as transcripts, small molecules, proteins, enzymes, or pH changes may interact with the host cell and cause differential transcription factor binding, methylation, or

chromatin availability, ultimately triggering differential expression of host genes and differential protein abundance (Nichols & Davenport, 2021) (Figure 20).

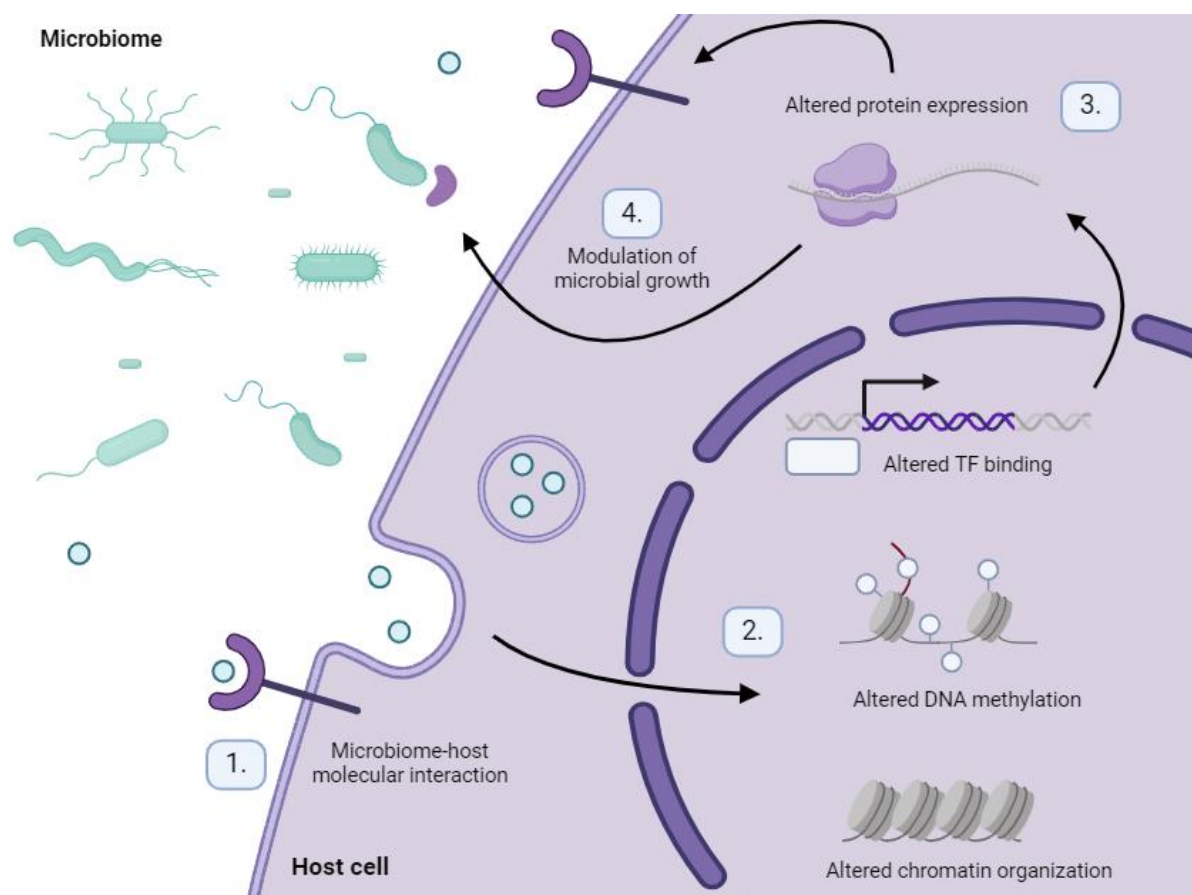


Figure 20. Potential mechanisms underlying host gene expression-microbiome associations. Created with BioRender.com. Adapted from Nichols R. et al., 2021, permission for figure adaptation granted by the Creative Commons license (<http://creativecommons.org/licenses/by/4.0/>.)

Nonetheless, these interactions are not necessarily pathogenic, as bacterial-derived molecules such as short-chain fatty acids (SCFAs), tryptophan, and bile acid metabolites are thought to help maintain gut, cardiovascular and systemic immune homeostasis; promoting the differentiation and function of immunosuppressive cells and also inhibiting inflammatory cells. Additionally, outer membrane vesicles (OMVs)

derived from bacteria have been demonstrated to transport virulence factors and short interfering RNAs to target host cells, disrupting host signaling pathways and modulating immune response (Charpentier et al., 2023). Thus, significant changes in microbiome composition may lead to dysregulated immune functions and associate with several conditions such as inflammatory bowel disease, tumors, metabolic syndromes, and cardiovascular diseases (Agus et al., 2021; Su et al., 2022; J. Wang et al., 2023; Z. Wang & Zhao, 2018).

Similar mechanisms have been proposed to occur in the endometrium. The commensal endometrial microbiome may engage in a symbiotic relationship with the endometrium and its local immune mediators, potentially facilitating implantation and modulating immune tolerance that is crucial for early pregnancy success. Thus, changes in microbiome composition could lead to dysregulation of inflammatory pathways in the endometrium, impacting endometrial homeostasis and contributing to infertility (Figure 21) (Odendaal et al., 2024). Interestingly, differential relative expression of tissue inflammatory mediators such as interferon (IFN)- γ and interleukins (IL)-6 and -1 β have been associated with endometrial dysbiosis in women with RIF and recurrent miscarriage (Cela et al., 2022; Liu et al., 2022), highlighting the significant interplay between the endometrial microbiota and local immune responses in the context of early pregnancy.

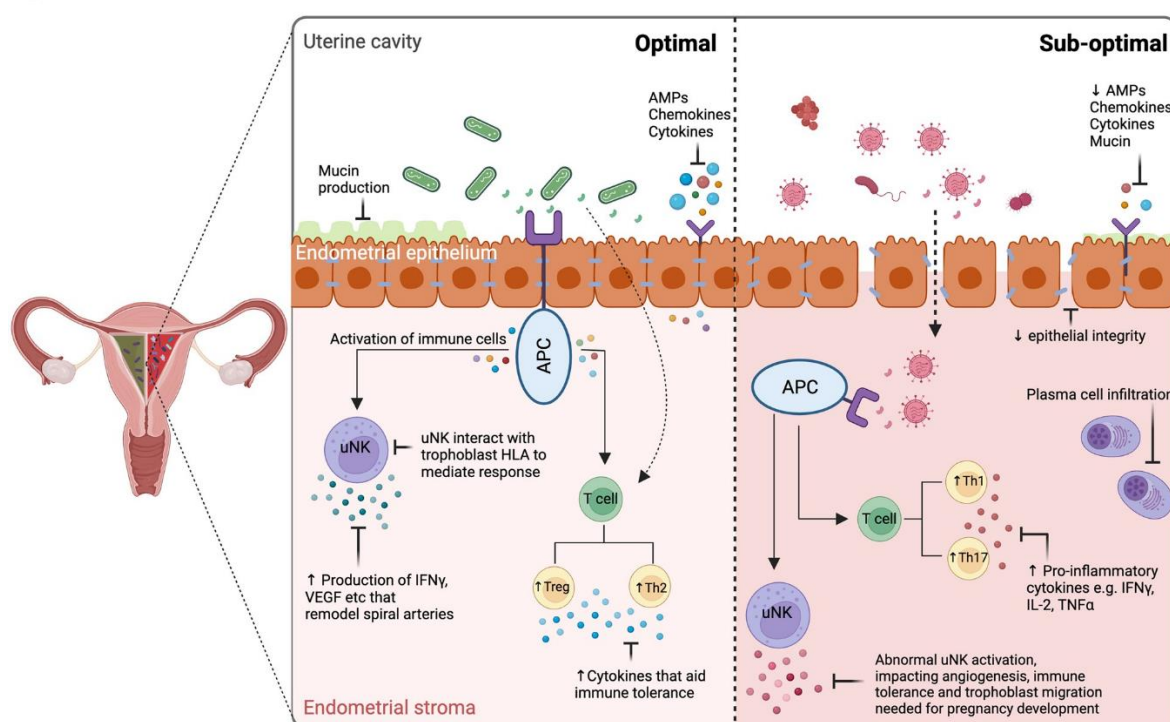


Figure 21. Potential mechanisms linking endometrial microbiota to early pregnancy loss. A shift towards a dysbiotic endometrial microbiota and increased pathogen colonization could dysregulate chemokine and cytokine production, resulting in loss of epithelial integrity, increased permeability, and bacterial translocation. Reprinted from Odendaal J. et al., 2024, permission for figure reproduction granted by the Creative Commons license (<http://creativecommons.org/licenses/by/4.0/>.)

Another example of host-microbe interaction happens through G protein-coupled receptors (GPCRs), which represent the largest family of receptors in the human genome. They are integral membrane proteins that play a pivotal role in cellular signaling by facilitating communication between the extracellular environment and the cell's physiology. Upon ligand binding, GPCRs undergo conformational changes that trigger the activation of G protein complexes to initiate downstream signaling cascades. These signaling pathways often involve the generation of second messengers, such as cyclic AMP (cAMP) or calcium ions (Ca²⁺), and culminate in a cellular response, which can range from changes in gene expression to alterations in cell survival or cell migration, depending on the specific GPCR and cell type involved (Figure 22) (Dorsam & Gutkind, 2007).

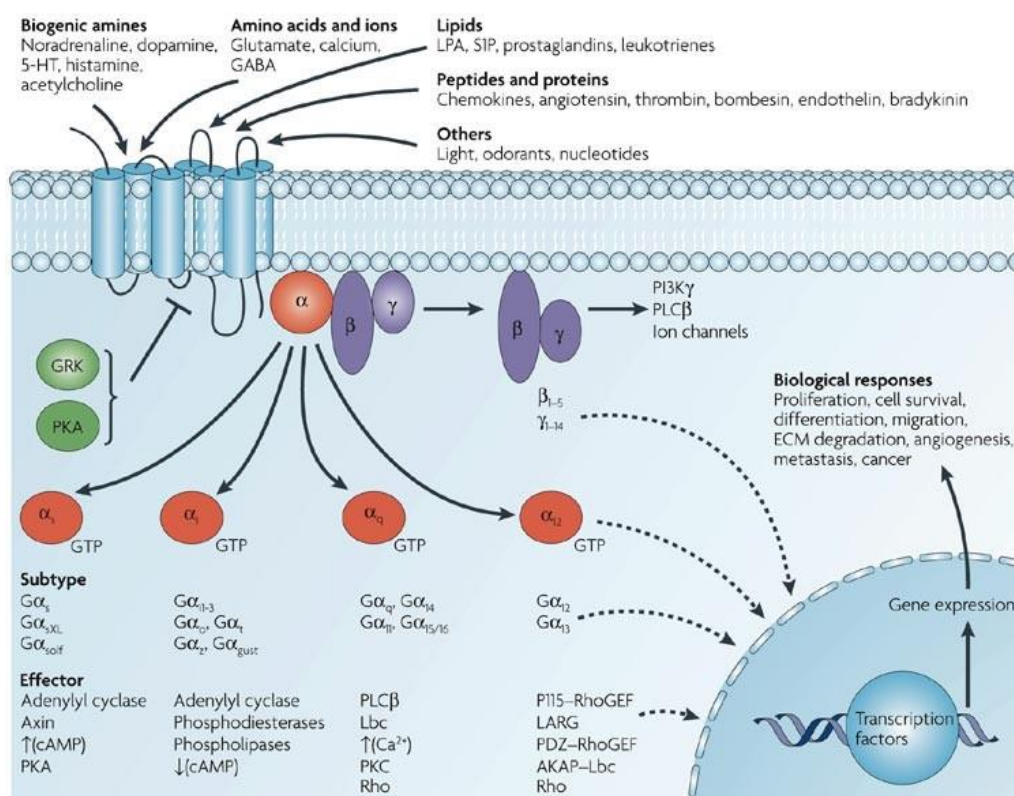


Figure 22. Diversity of G-protein-coupled receptor signaling. Various ligands use GPCRs to stimulate membrane, cytoplasmic and nuclear targets. Reprinted from Dorsam & Gutkind, 2007 with permission from Springer Nature.

Commensal bacteria were also demonstrated to express metabolites that mimic human signaling molecules, being able to interact with several host proteins including GPCRs (Figure 23) (Cohen et al., 2017; Colosimo et al., 2019). As an example, bacterial ligands may trigger several human receptors such as hydroxycarboxylic acid receptors, histamine receptors, and even neurotransmitter receptors (Colosimo et al., 2019). Additionally, novel high throughput methods have shown hundreds of bacterial strains isolated from the human gut, lung, mouth, skin, and vagina are able to trigger GPCR activation, revealing a complex microbiome-GPCR interaction network that may happen in different body sites (Chen et al., 2023). Importantly, by performing a cross-tissue human microbiome-GPCR interaction atlas, they observed that 15 strains of *G. vaginalis* and *L. iners* from the vaginal microbiome are able to non-specifically

activate numerous GPCRs, suggesting that these strains may produce compounds that act as pan activators of the host transcriptional program leading to human biological responses to bacterial-secreted stimuli (Chen et al., 2023). Ultimately, these results highlight the significant impact bacterial metabolites could have on virtually all cells and tissues from the human body, demonstrating the importance of further studies regarding whole bacterial genomes and microbiome dynamics.

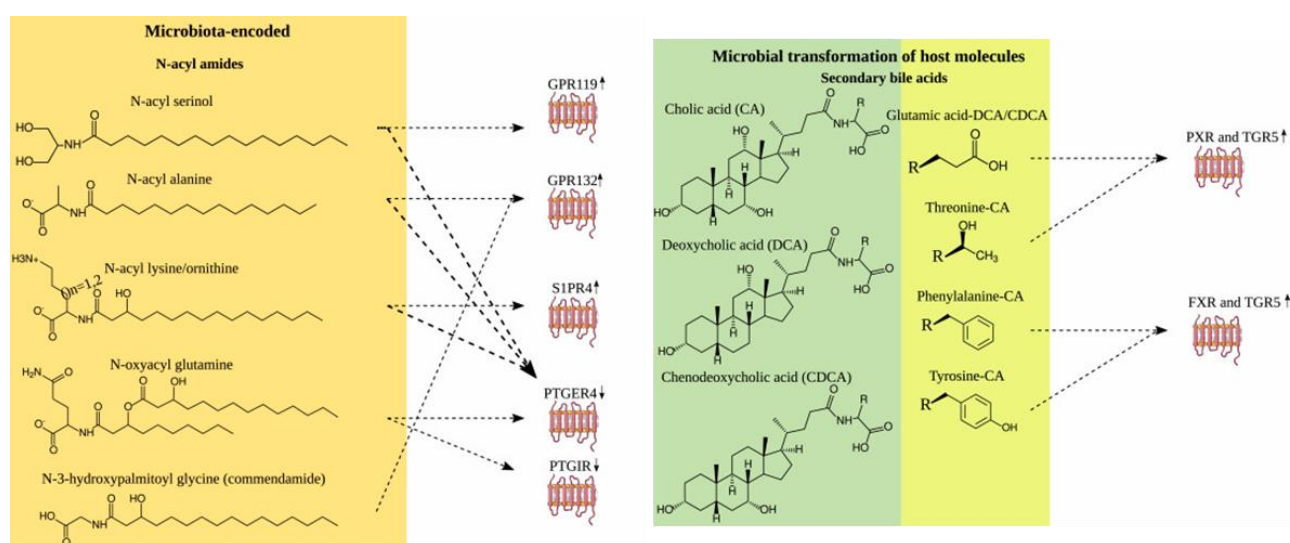


Figure 23. Overview of some microbial-derived metabolites known to interact with human receptors. Two broad categories of microbial metabolites include microbiota-encoded molecules and human molecules transformed by microbes. Up or down arrows to the right of the human receptors indicate microbial metabolites with agonist and antagonist activities, respectively. Adapted from (Aleti et al., 2023), permission for figure reproduction granted by the Creative Commons license (<http://creativecommons.org/licenses/by/4.0/>.)

Importantly, dysregulated signaling of several GPCRs has been demonstrated to significantly impact reproductive health not only by modulating menstrual cycle and decidualization, but also by playing a toll in the onset of gynecological diseases such as endometriosis, adenomyosis, and preeclampsia (Table 2).

Table 2. Overview of GPCRs previously associated with gynecological processes and diseases.

GPCR signaling	Gynecological condition/process	Reference
Relaxin/insulin-like family peptide receptor 1 (RXFP1)	Preeclampsia	(Conrad, 2016)
Prokineticin 2 receptor (PKR1)	Decidualization and embryo implantation	(J. Zhai et al., 2022)
Secretin/adhesion (Class B) GPCRs	Placental development and preeclampsia	(Yin et al., 2022a)
Neuropeptide S receptor 1 (NPSR1)	Endometriosis	(Tapmeier et al., 2021)
Stanniocalcin-1 (ligand of calcium-sensing receptor, CASR)	Endometriosis	(Aghajanova et al., 2016)
Mu opioid receptor (MOR)	Menstrual cycle	(Totorikaguena et al., 2017)
Lysophosphatidic acid receptor (LPAR3)	Decidualization	(Aikawa et al., 2017)
Calcitonin receptor-like receptor (CRLR)	Adenomyosis	(Xu et al., 2021)
Adrenergic β 2 receptor (ADRB2)		
Tachykinin receptor 1 (NK1R)		

Abbreviation: **GPCR**, G protein-coupled receptor.

Thus, considering the growing evidence that the microbiota modulates host cell metabolism via GPCR signaling, and GPCR dysregulation associates with reproductive health, it is feasible to speculate that microbial modulation of these receptors might represent a significant mechanism for microbiome-associated gynecological diseases, and further studies should be able to unravel these features.

1.4.4 Host cell response to dysbiotic microbiota and infection

Regardless of all the molecular mechanisms by which microbes modulate target cell metabolism, the host also plays a crucial role in shaping and maintaining the microbiota composition. In fact, there is a complex interplay between the host and

microbiota that involves various mechanisms and influences the abundance and diversity of microbial communities.

For instance, host-induced compartmentalization of the gut microbiota via dendritic cells, secretory IgA antibodies, and antimicrobial peptides (AMPs) have been described to help maintain the mucosal barrier function (Zheng et al., 2020). Additionally, pattern recognition receptors such as Toll-like receptors help in the defense against pathogens and regulate the abundance of commensal microbes, maintaining tissue integrity (Rakoff-Nahoum et al., 2004). For instance, TLR activation triggers several host reactions: secretion of AMPs, IgA, and cytoprotective factors such as trefoil factor 3 (TFF3); tightening of intercellular junctions between epithelial cells; and promotion of cellular immune response by releasing inflammatory cytokines (Figure 24) (Burgueño & Abreu, 2020).

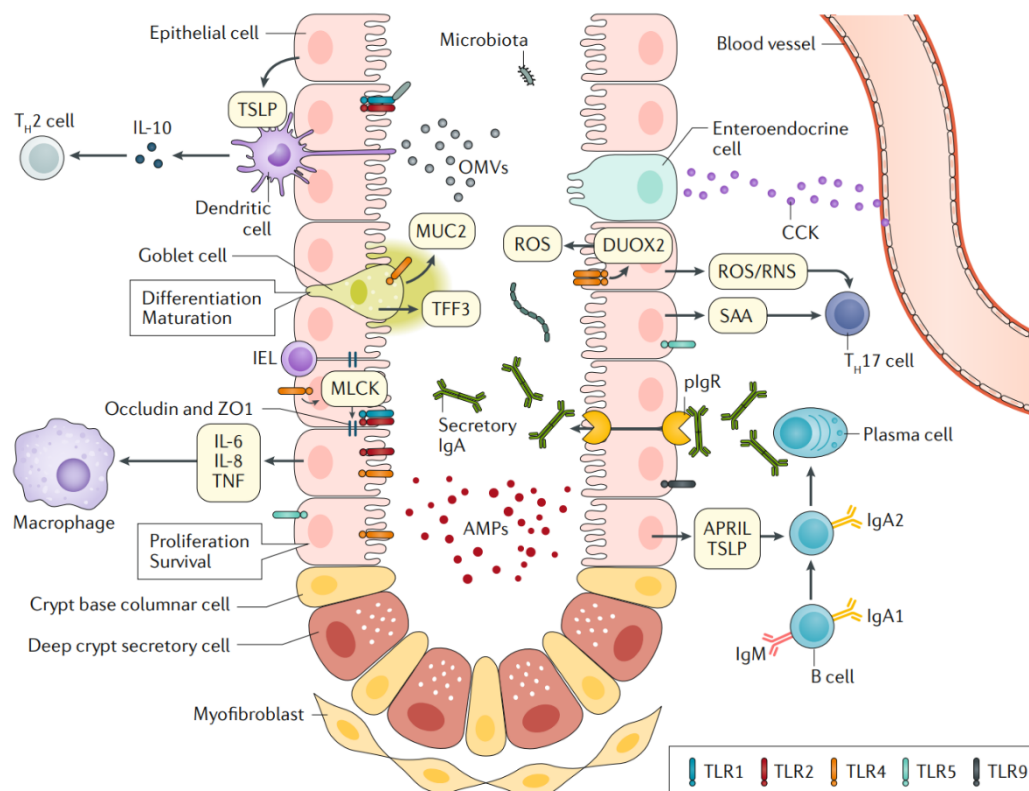


Figure 24. Scheme of host mechanisms for microbiota maintenance and regulation in the gut. Reprinted from Burgueño & Abreu, 2020 with permission from Springer Nature.

Conversely, the immune system has also been demonstrated to play a significant role in uterine and reproductive health. For instance, the maternal-embryo interface displays a highly active and complex immunological environment, as a genetically distinct fetus must be supported for the required gestational period. Nonetheless, besides being responsible for tolerance towards the embryo, the uterine environment must keep up the defenses against microbial pathogens (Semmes & Coyne, 2022). Additionally, a tight immune regulation is also needed for embryo implantation and pregnancy maintenance, as dysregulation in cytokine expression and immune cell activation has been associated with implantation failure and pregnancy loss (Pantos et al., 2022; Yang et al., 2022). Lastly, immunological changes have been reported in important gynecological conditions, such as preeclampsia, adenomyosis, and endometriosis (Bourdon, Santulli, Jeljeli, et al., 2021; Deer et al., 2023; Králíčková et al., 2018).

Although the mechanisms underlying innate and adaptive immune response in pregnancy and reproductive health are complex and still need to be fully elucidated, it is feasible to speculate that variations in immune regulations within the uterine cavity might have a strong impact in the endometrial microbiota (Gholiof et al., 2022). Nonetheless, genetics account for 20-40% of immune variation, meaning that genetic variations may also influence how individuals respond to fluctuations in their microbiota (Häder et al., 2023; Liston et al., 2016). Thus, deep understanding of the interplay between host genetics and immune response with the microbiota is another essential topic to fully elucidate the impact of microbial dynamics in reproductive health, and could help explain why patients with the same microbiota composition may bear distinct reproductive outcomes. Ultimately, the genetic background may

significantly contribute to individual variations in microbial composition, immune responses, and overall health.

1.4.5 Applications in reproductive medicine

Due to the increasing comprehension of bacteria influence on fertility and various gynecological conditions, modulating the microbiota composition via antibiotics or probiotics use has gained strength in reproductive medicine. Building on this knowledge, microbe-derived metabolites, virulence factors, and ARGs further underscore the significance of understanding the microbiome in reproductive medicine. Specifically, comprehending the impact of bacterial-derived factors on the FGT is essential to unravel the mechanisms behind the onset of gynecological conditions and infertility, as they may modulate cell host metabolism, colonization success, and therapy efficiency. Additionally, building knowledge on the host factors that modulate microbiota dynamics should also promote important insights, given the strong impact host genetics and immune response poses on microbiome.

Understanding the interactions between bacterial and host factors and their specific modes of action can provide valuable insights into the intricate interplay between microbial communities and reproductive health. Ultimately, this knowledge will pave the way for the development of innovative treatments that go beyond traditional antibiotic and probiotic approaches, offering novel therapeutic strategies to address reproductive conditions and enhance fertility outcomes.

2. BACKGROUND AND HYPOTHESIS

2.1 BACKGROUND

Results gathered throughout the years have consistently demonstrated the existence of an endogenous endometrial microbiota, challenging the dogma of the sterile uterus. Building on this knowledge, previous research from our group used 16S rRNA gene sequencing to analyze endometrial fluid and biopsy samples before embryo transfer in a cohort of 342 infertile patients undergoing ART. This multicenter observational study showed higher abundance of *Lactobacillus* associates with reproductive success, while a dysbiotic profile (higher abundance of pathogens such as *Atopobium* spp., *G. vaginalis*, and *Bifidobacterium* spp., among others) associates with decreased implantation, pregnancy, ongoing pregnancy, and live birth rates per transfer (Moreno et al., 2022). Although these findings served for a novel personalized diagnosis of endometrial microbiota in the context of reproductive medicine, little information has been reported regarding the bacterial/host mechanisms leading to such associations.

2.2 HYPOTHESIS

Considering recent findings regarding the impact of the endometrial microbiota in reproductive medicine, we hypothesize that bacteria within the uterine cavity may significantly contribute to cell host metabolism and the overall uterine environment. Thus, we propose that the uterine microbiome, in normal and dysbiotic conditions, contributes via a molecular level to uterine pathophysiology that creates a niche more beneficial or detrimental to embryo implantation and a healthy pregnancy.

3. OBJECTIVES

The general aim of this project was to characterize the microbiome composition within the uterine cavity and its functional impact on reproductive outcomes of women undergoing ART.

To accomplish this, we set the following specific objectives:

- I. To characterize the main bacterial profiles of the uterine microbiome, at the species level, associated to reproductive success vs. failure using whole metagenomic sequencing (WMS) from endometrial fluid (EF)
- II. To identify bacterial functions, as well as virulence factors and antibiotic resistance genes, enriched in reproductive failure and/or success
- III. To evaluate the impact of normal vs. dysbiotic uterine microbiome on endometrial gene expression profiles by metatranscriptomic (RNAseq) analyses of the endometrium

4. METHODS

4.1 PATIENTS, SAMPLES, AND ETHICS

Samples for this study were selected from a prospective multicenter observational research analyzing the association between the endometrial microbiota composition and reproductive outcomes in patients undergoing ART previously published by our group (Moreno et al., 2022).

Out of 342 patients enrolled in the previous study, 79 women were selected after fulfilling the following criteria:

- I. Women with known reproductive outcomes after a personalized embryo transfer in a receptive endometrium
- II. Women with sufficient sample volume to perform paired WMS in DNA obtained from EF samples and MTX analysis in total RNA obtained from EB samples
- III. Women with a detectable microbiome in EF samples analyzed by 16S rRNA-seq

The corresponding local Ethics Committees gave ethical approval according to the protocol with reference IGX1-MIC-CS-17-05 by Western Institutional Review Board (IVF-Florida, USA, study code 1176556, July 17, 2017; RMA Connecticut, USA, study code 1177331, July 31, 2017; Dominion Fertility, USA, study code 1176555, July 22, 2017; Missouri Center for Reproductive Medicine, USA, study code 1179405, October 10, 2017; Pacific Centre for Reproductive Medicine, Canada, study code 1179667, December 13, 2017). All participants provided written informed consent and were

classified into different groups according to the following criteria (Zegers-Hochschild et al., 2017):

- I. **Reproductive success:** refers specifically to the occurrence of a live birth. Achieving reproductive success underscores the successful progression of all biological processes associated with human conception (fertilization, implantation, embryonic development, and gestation), ultimately resulting in the birth of a live infant.
- II. **Reproductive failure:** comprises a spectrum of outcomes that deviate the pregnancy from achieving a successful live birth, including:
 - Biochemical pregnancy** is an early form of reproductive failure, characterized by the early detection of beta human chorionic gonadotropin (β -hCG) in serum or urine, without subsequent embryonic development.
 - Clinical miscarriage** involves the loss of an intra-uterine pregnancy prior to 22 completed weeks of gestational age.
 - Ectopic pregnancy** occurs when the fertilized egg implants outside the uterus, diagnosed by ultrasound, surgical visualization or histopathology.

Finally, the absence of pregnancy itself can be viewed as a form of reproductive challenge, prompting exploration into potential factors influencing conception. Thus, “**no pregnancy**” applies to those patients who did not have positive β -hCG results, neither had a visible gestational sac after ultrasound examination following pET.

After receiving pET at the time of maximum endometrial receptivity following endometrial receptivity analysis, 27 women achieved a successful outcome in terms of live birth (LB, 34,2%), while 52 presented reproductive failure: 38 did not become pregnant (NP, 48,1%), 4 with a biochemical pregnancy (BP, 5%), 9 had a clinical

miscarriage (CM, 11,4%), and 1 developed an ectopic pregnancy (EP, 1,3%) (Figure 25).

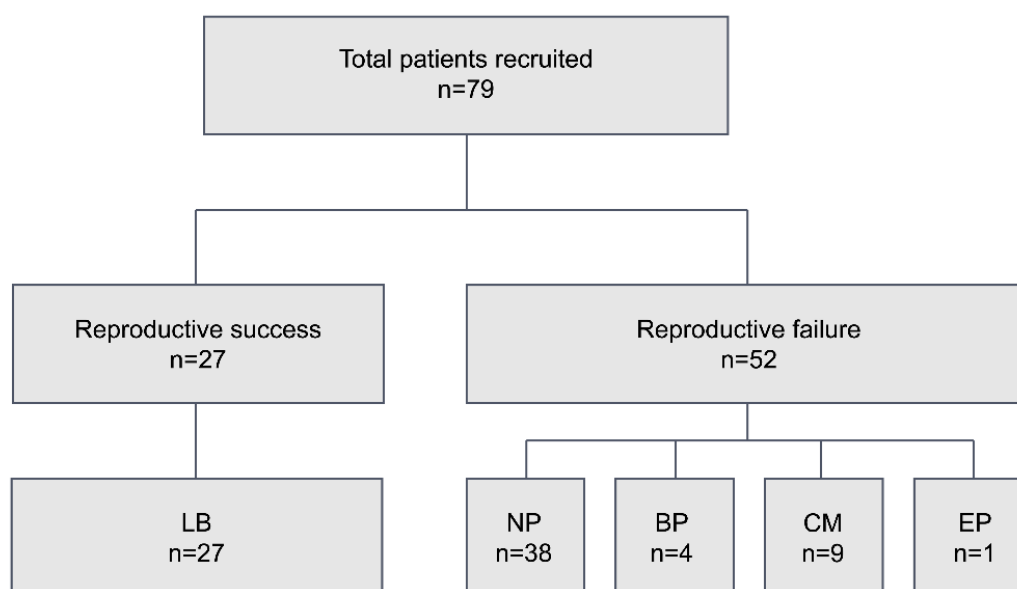


Figure 25. Patients allocated in the study classified according to their reproductive outcomes after assisted reproductive technology. Abbreviations: **BP**, biochemical pregnancy; **CM**, clinical miscarriage; **EP**, ectopic pregnancy; **LB**, live birth; **NP**, no pregnancy.

After analyzing the demographics (Table 3) to search for potential confounding factors, no bias toward the clinical outcome for these study groups were observed.

Table 3. Sociodemographic and clinical characteristics of study participants.

	Total (n=79)	Live birth (n=27)	No pregnancy (n=38)	Biochemical pregnancy (n=4)	Clinical miscarriage (n=9)	Ectopic pregnancy (n=1)
Age	34.8 (5.0)	34.0 (4.3)	35.7 (5.6)	37.5 (1.3)	32.8 (4.7)	29.0 (NA)
BMI	23.4 (3.3)	23.7 (3.3)	23.1 (3.2)	23.5 (1.9)	24.6 (3.9)	18.6 (NA)
Ethnicity						
African	1 (1.3%)	1 (3.7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Caucasian	43 (54.4%)	15 (55.6%)	24 (63.2%)	1 (25%)	3 (33.3%)	0 (0%)
East Asian	14 (17.7%)	2 (7.4%)	8 (21.1%)	1 (25%)	2 (22.2%)	1 (100%)
Hispanic	8 (10.1%)	2 (7.4%)	4 (10.5%)	0 (0%)	2 (22.2%)	0 (0%)
Multiracial	1 (1.3%)	1 (3.7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
South Asian	3 (3.8%)	2 (7.4%)	0 (0%)	1 (25%)	0 (0%)	0 (0%)
Other	9 (11.4%)	4 (14.8%)	2 (5.3%)	1 (25%)	2 (22.2%)	0 (0%)

	Total (n=79)	Live birth (n=27)	No pregnancy (n=38)	Biochemical pregnancy (n=4)	Clinical miscarriage (n=9)	Ectopic pregnancy (n=1)
Previous ectopic pregnancies, #	0.1 (0.4)	0.2 (0.5)	0.1 (0.4)	0.0 (0.0)	0.0 (0.0)	1.0 (NA)
Previous stillbirth, #	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (NA)
Previous live birth, #	0.3 (0.6)	0.4 (0.6)	0.2 (0.4)	0.0 (0.0)	0.4 (1.0)	0.0 (NA)
Implantation failure	2.0 (1.5)	2.0 (1.4)	2.4 (1.6)	0.5 (0.6)	1.2 (1.3)	3.0 (NA)
ART Indication						
AMA	26 (32.9%)	8 (29.6%)	13 (34.2%)	3 (75%)	2 (22.2%)	0 (0%)
Male factor	29 (36.7%)	10 (37%)	11 (28.9%)	2 (50%)	6 (66.7%)	0 (0%)
RPL	4 (5.1%)	2 (7.4%)	0 (0%)	0 (0%)	2 (22.2%)	0 (0%)
RIF	23 (29.1%)	6 (22.2%)	15 (39.5%)	0 (0%)	1 (11.1%)	1 (100%)
Tubal factor	7 (8.9%)	3 (11.1%)	3 (7.9%)	0 (0%)	1 (11.1%)	0 (0%)
Uterine factor	1 (1.3%)	0 (0%)	0 (0%)	0 (0%)	1 (11.1%)	0 (0%)
Ovarian pathology	23 (29.1%)	7 (25.9%)	10 (26.3%)	4 (100%)	2 (22.2%)	0 (0%)
Endometriosis	4 (5.1%)	3 (11.1%)	1 (2.6%)	0 (0%)	0 (0%)	0 (0%)
Genetic	3 (3.8%)	0 (0%)	2 (5.3%)	0 (0%)	1 (11.1%)	0 (0%)
Infertility	11 (13.9%)	5 (18.5%)	3 (7.9%)	1 (25%)	2 (22.2%)	0 (0%)
No indication	1 (1.3%)	1 (3.7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
>1 indication	38 (48.1%)	13 (48.1%)	16 (42.1%)	4 (100%)	5 (55.6%)	0 (0%)
FSH mIU/ml*	6.8 (3.2)	6.0 (1.7)	6.3 (4.8)	10.0 (2.3)	7.3 (3.0)	NA
AMH ng/ml*	2.8 (2.1)	2.3 (2.1)	2.9 (2.1)	2.4 (2.1)	4.3 (2.4)	NA

Data are expressed as mean (SD) or n (%). Non-parametric Kruskal-Wallis test compared numerical variables in each group. Wilcoxon rank-sum test compared quantity variables between groups. Chi-Square and Fisher exact tests compared categorical variables. Bonferroni was used for multiple comparison testing. Abbreviations: **AMA**, advanced maternal age; **AMH**, anti-Müllerian hormone; **ART**, assisted reproductive treatment; **BMI**, body-mass index (kg/m²); **FSH**, follicle stimulating hormone; **RIF**, recurrent implantation failure; **RPL**, recurrent pregnancy loss. * Hormone levels not measured in women receiving embryo transfer with donated oocytes.

Additionally, the same assessment was performed for clinical characteristics of patients, and no bias were found (Table 4).

Table 4. Clinical variables in patients with different reproductive outcomes.

	Total (n=79)	Live birth (n=27)	No pregnancy (n=38)	Biochemical pregnancy (n=4)	Clinical miscarriage (n=9)	Ectopic pregnancy (n=1)
Treatment						
ICSI frozen	60 (75.9%)	20 (74.1%)	28 (73.7%)	2 (50%)	9 (100%)	1 (100%)
IVF frozen	3 (3.8%)	1 (3.7%)	2 (5.3%)	0 (0%)	0 (0%)	0 (0%)
IVF/ICSI frozen	16 (20.3%)	6 (22.2%)	8 (21.1%)	2 (50%)	0 (0%)	0 (0%)
Stimulation protocol						
Long	19 (24.1%)	9 (33.3%)	6 (15.8%)	2 (50%)	2 (22.2%)	0 (0%)
Short	49 (62%)	17 (63%)	22 (57.9%)	2 (50%)	7 (77.8%)	1 (100%)
N/A	11 (13.9%)	1 (3.7%)	10 (26.3%)	0 (0%)	0 (0%)	0 (0%)
Egg Source						
Own	62 (78.5%)	24 (88.9%)	26 (68.4%)	4 (100%)	7 (77.8%)	1 (100%)
Donated	17 (21.5%)	3 (11.1%)	12 (31.6%)	0 (0%)	2 (22.2%)	0 (0%)
Ovulation induction						
Agonist	19 (24.1%)	9 (33.3%)	6 (15.8%)	2 (50%)	2 (22.2%)	0 (0%)
Double Triggering	12 (15.2%)	5 (18.5%)	6 (15.8%)	0 (0%)	1 (11.1%)	0 (0%)
HCG	37 (46.8%)	12 (44.4%)	16 (42.1%)	2 (50%)	6 (66.7%)	1 (100%)
N/A	11 (13.9%)	1 (3.7%)	10 (26.3%)	0 (0%)	0 (0%)	0 (0%)
Antral follicles, #	16.1 (8.0)	17.3 (8.2)	14.6 (7.5)	15.2 (10.9)	17.1 (9.0)	19.0 (NA)
MII oocytes, #	11.3 (5.8)	12.7 (6.1)	9.8 (5.1)	8.8 (6.5)	12.2 (6.4)	20.0 (NA)
Total hours P4 in sample collection cycle	122.1 (8.2)	124.0 (9.9)	120.4 (7.6)	121.6 (2.8)	124.1 (6.1)	120.0 (NA)
Total hours P4 recommended for ET	127.9 (13.1)	128.9 (13.9)	126.7 (12.8)	118.8 (9.9)	133.4 (12.6)	132.0 (NA)
Total hours P4 in ET cycle	127.7 (13.1)	128.9 (14.4)	126.3 (12.3)	118.7 (9.9)	133.7 (13.3)	128.2 (NA)
Antibiotic intake in ET						
YES	10 (12.7%)	4 (14.8%)	3 (7.9%)	1 (25%)	2 (22.2%)	0 (0%)
NO	69 (87.3%)	23 (85.2%)	35 (92.1%)	3 (75%)	7 (77.8%)	1 (100%)
Probiotics before ET						
YES	8 (10.1%)	2 (7.4%)	3 (7.9%)	1 (25%)	2 (22.2%)	0 (0%)
NO	63 (79.7%)	22 (81.5%)	32 (84.2%)	2 (50%)	6 (66.7%)	1 (100%)
UNKNOWN	8 (10.1%)	3 (11.1%)	3 (7.9%)	1 (25%)	1 (11.1%)	0 (0%)
Days between sample collection and ET	87.3 (73.2)	78.7 (56.9)	100.3 (91.8)	58.8 (18.2)	72.7 (30.2)	66.0 (NA)
Embryos transferred	1.2 (0.4)	1.3 (0.5)	1.2 (0.4)	1.0 (0.0)	1.1 (0.3)	1.0 (NA)
Embryo stage						
Cavitated blastocyst	11 (13.9%)	1 (3.7%)	6 (15.8%)	1 (25%)	2 (22.2%)	1 (100%)
Early blastocyst	5 (6.3%)	0 (0%)	5 (13.2%)	0 (0%)	0 (0%)	0 (0%)
Expanded blastocyst	19 (24.1%)	10 (37%)	7 (18.4%)	0 (0%)	2 (22.2%)	0 (0%)
Hatched blastocyst	2 (2.5%)	2 (7.4%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Hatching blastocyst	42 (53.2%)	14 (51.9%)	20 (52.6%)	3 (75%)	5 (55.6%)	0 (0%)

	Total (n=79)	Live birth (n=27)	No pregnancy (n=38)	Biochemical pregnancy (n=4)	Clinical miscarriage (n=9)	Ectopic pregnancy (n=1)
Embryo quality						
AA	10 (12.7%)	2 (7.4%)	7 (18.4%)	0 (0%)	1 (11.1%)	0 (0%)
AB	10 (12.7%)	4 (14.8%)	5 (13.2%)	0 (0%)	1 (11.1%)	0 (0%)
BA	4 (5.1%)	0 (0%)	2 (5.3%)	1 (25%)	1 (11.1%)	0 (0%)
BB	24 (30.4%)	10 (37%)	8 (21.1%)	2 (50%)	4 (44.4%)	0 (0%)
BC	20 (25.3%)	8 (29.6%)	8 (21.1%)	1 (25%)	2 (22.2%)	1 (100%)
CA	1 (1.3%)	1 (3.7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
CC	9 (11.4%)	2 (7.4%)	7 (18.4%)	0 (0%)	0 (0%)	0 (0%)
N/A	1 (1.3%)	0 (0%)	1 (2.6%)	0 (0%)	0 (0%)	0 (0%)

Data are expressed as mean (SD) or n (%). Non-parametric Kruskal-Wallis test compared numerical variables in each group. Wilcoxon rank-sum test compared quantity variables between groups. Chi-Square and Fisher exact tests compared categorical variables. Bonferroni was used for multiple comparison testing. Abbreviations: **ET**, embryo transfer; **HCG**, human chorionic gonadotropin; **ICSI**, intracytoplasmic sperm injection; **IVF**, in vitro fertilization; **MII**, metaphase II; **NA**, not applicable; **P4**, progesterone.

4.2 GENOMIC DNA AND RNA ISOLATION

Total DNA was isolated from EF samples by performing a pre-digestion step at 37°C for 30 min with 25 µg/µL lysozyme, 0.12 U/µL lysostaphin, 0.4 U/µL mutanolysin, and 1.8% Triton X-100 to degrade bacterial cell walls. DNA was then extracted with the QIAamp DNA Blood Mini kit (Qiagen) following the manufacturer's instructions. Finally, DNA was eluted with 35 µL of nuclease-free water and quantified using Tape Station (Agilent Technologies).

Total RNA was obtained from approximately 25 mg of EB samples using the RNeasy Mini Kit (Qiagen). EB samples were mechanically disaggregated and homogenized using the bead-beating method with 5 mm diameter stainless steel beads (Qiagen) and a mix of RLT buffer and β-mercaptoethanol (Sigma Aldrich) in a TissueLyser LT (Qiagen) for 5 min at 50 Hz. RNA was then purified by adding ethanol

to the lysate and subsequent passage of the material through the RNeasy Mini spin columns (Qiagen). The resulting RNA was quantified using the Tape Station (Agilent Technologies).

4.3 WHOLE METAGENOME SEQUENCING

Bacterial taxonomical and functional composition in the EF samples was assessed by WMS using the Illumina DNA Prep kit (Illumina) following the manufacturer's instructions. Total DNA samples were normalized to 30 ng for library preparation. Tagmentation of genomic DNA was performed using Bead-linked transposomes, followed by clean-up. Limited-cycle PCR was performed to amplify the tagmented DNA and add the indexes/adapters required for cluster generation. Small library fragments were filtered out using Agencourt AMPure XP beads (Beckman Coulter, Inc.), and library quality was assessed using Tape Station (Agilent Technologies). The pooled libraries were paired-end sequenced (2 x 150 bp) on the Illumina NovaSeq 6000 platform at an expected average sequencing depth of 150M reads/sample.

4.4 METATRANSCRIPTOMICS

The Illumina Stranded Total RNA with RiboZero Plus Kit (Illumina) was used to prepare MTX libraries. Total RNA samples were normalized to 100 ng and used for ribosomal RNA depletion. RNA was then fragmented and annealed with random hexamers, which prime the sample for cDNA synthesis. Once the cDNA was synthesized, anchors were ligated to the end of the double-stranded cDNA to prepare for dual indexing. The library was amplified after cleaning up the fragments with Agencourt AMPure XP magnetic beads (Beckman Coulter, Inc.), and RNA UD indexes

(Illumina) were added. PCR was performed to selectively amplify the anchor-ligated DNA fragments and add the indexes and primer sequences required for cluster generation. After cleaning up the sample using Agencourt AMPure XP beads (Beckman Coulter, Inc.), the library quality was assessed using the Tape Station (Agilent Technologies). The pooled libraries were paired-end sequenced (2 x 100 bp) on the Illumina NovaSeq 6000 platform at an expected average sequencing depth of 100M reads/sample.

4.5 BIOINFORMATIC SEQUENCING ANALYSIS: METAGENOMICS

4.5.1 Sequencing-based microbiome analysis

Reads were quality-trimmed and length-filtered using PRINSEQ (Schmieder & Edwards, 2011). Paired-end reads were merged using the Fast Length Adjustment of Short reads (FLASH) software tool (Magoč & Salzberg, 2011) and, finally, host-reads were removed using Burrows-Wheeler Aligner (BWA) mapped against human genome reference (H. Li & Durbin, 2010). Bacterial genera not colonizing humans or consistently associated with kitome contaminants were also removed from the analysis (Barton et al., 2006; Glassing et al., 2016; Grahn et al., 2003; Hashimoto & Kyono, 2019; Kitaya et al., 2019; Kyono, Hashimoto, Nagai, et al., 2018; Lauder et al., 2016; Laurence et al., 2014; Stinson et al., 2019; Tanner et al., 1998; Weyrich et al., 2019). Additionally, although a few studies have detected *Streptomyces* spp. in clinical samples (Gras et al., 2022), microbes belonging to this genus are mostly found in soil, water and plants, and thus can be considered as contaminants and have been also excluded from analyses.

Functional and taxonomical profiles were obtained using Diamond alignment analysis (Buchfink et al., 2021) using a non-redundant NCBI Reference Sequence

database (Pruitt, 2005). This method annotates read-by-read, retrieving the BLAST alignment result of each read in the database. Taxonomical and functional abundances were estimated by parsing BLAST output results. Finally, taxonomical and functional outputs were processed using R statistical software (R Core Team, 2018) for statistical description and graphical representation of the sample's taxonomical and functional profile.

4.5.2 Community state types (CSTs)

Community state types (CSTs) were obtained using hierarchical clustering over Euclidean distance between samples' relative normalized abundance taxa composition. R statistical software (R Core Team, 2018) was used for statistical and graphical representation.

4.5.3 Correlation network analysis

Co-occurrence network analysis was done applying SparCC (Friedman & Alm, 2012) correlation algorithm to merge raw counts of taxonomical and functional annotation from metagenomic results. The implementation of algorithm was done using SpiecEasi R library (Kurtz et al., 2015). Taxonomic networks were visualized with the igraph library (Csardi & Nepusz, 2006) in R programming environment for statistical analysis (R Core Team, 2018), considering taxa with correlation coefficients higher than 0.7. Taxa-function associations were visualized with the igraph and corrplot library (T. Wei et al., 2021). Negative correlations were represented as red, while positive correlations are in blue.

4.5.4 Virulence factors and antibiotic resistance analysis

Virulence Factor Database (VFDB) (L. Chen, 2005) was used together with Diamond sequence aligner for protein and translated RNA searches (Buchfink et al., 2021) in the case of virulence factors. Profiles of antibiotic resistance were obtained from samples by mapping with the Comprehensive Antibiotic Resistance Database (CARD) and Resistance Gene Identifier (RGI) alignment bioinformatic tools (Alcock et al., 2023). In both cases, differential abundances between outcomes were reported using DESeq2 library in the R statistical programming environment (R Core Team, 2018).

4.6 BIOINFORMATIC SEQUENCING ANALYSIS: METATRANSCRIPTOMICS

4.6.1 Sequencing-based RNA analysis

Reads generated by the Illumina sequencing platform were quality filtered and annotated using the SAMSA2 complete metatranscriptome analysis pipeline (Westreich et al., 2018). To align the metatranscriptomic reads, all reads were first mapped against the human genome using BWA (H. Li & Durbin, 2009). Subsequently, SAMtools was employed for processing the alignment files (.bam) and reads marked as unmapped were retrieved (H. Li et al., 2009). Ribosomal fractions of reads were manually removed from the SAMSA2 output. Finally, taxonomical and functional outputs from SAMSA2 were processed using R statistical software (R Core Team, 2018) for statistical description and graphical representation of the sample's taxonomical and functional profile.

4.7 STATISTICAL ANALYSIS

Fisher's exact test and Individual proportion test were performed using R statistical software for correlation assessment between CSTs and reproductive outcomes. Significant values were considered when $p < 0.05$.

All WGS and MTX datasets were analyzed for differentially abundant taxa and functions using DESeq2 (Love et al., 2014) and LEfSe (Segata et al., 2011). For human transcriptomics analysis, the edgeR R library was used for differential gene expression detection (Robinson et al., 2010), and the clusterProfiler R library was used for gene set enrichment with gene ontology (GO) terms (T. Wu et al., 2021). Significant values were considered when $\log_2\text{FoldChange}$ (FC) was ≥ 2.0 or ≤ -2.0 and their adjusted p-values (p-adj) or false discovery rate (FDR) values were < 0.05 .

5. RESULTS

5.1 CHARACTERIZATION OF ENDOMETRIAL BACTERIAL COMMUNITIES

After DNA extraction, we obtained an average of 171.09 ng/ μ L of DNA per sample (range 0.1 ng/ μ L – 1,009.2 ng/ μ L). Library preparation yielded an average of 2.36 ng/ μ L (range 0.1 ng/ μ L – 16.7 ng/ μ L). Importantly, out of 79 paired samples, 3 EF samples (1 LB and 2 NP) did not produce sequencing libraries. As these samples did not have sufficient DNA for library repetition, they were discarded from subsequent analysis; thus leaving 76 EF valid samples distributed into reproductive success (LB) or failure, which was further subdivided into no pregnancy, biochemical pregnancy, clinical miscarriage, and ectopic pregnancy (Figure 26). Notably, the study sample included a single patient from the EP group. While this patient was included in the reproductive failure group (taken as a whole), we do not describe this patient/EP separately as an outcome.

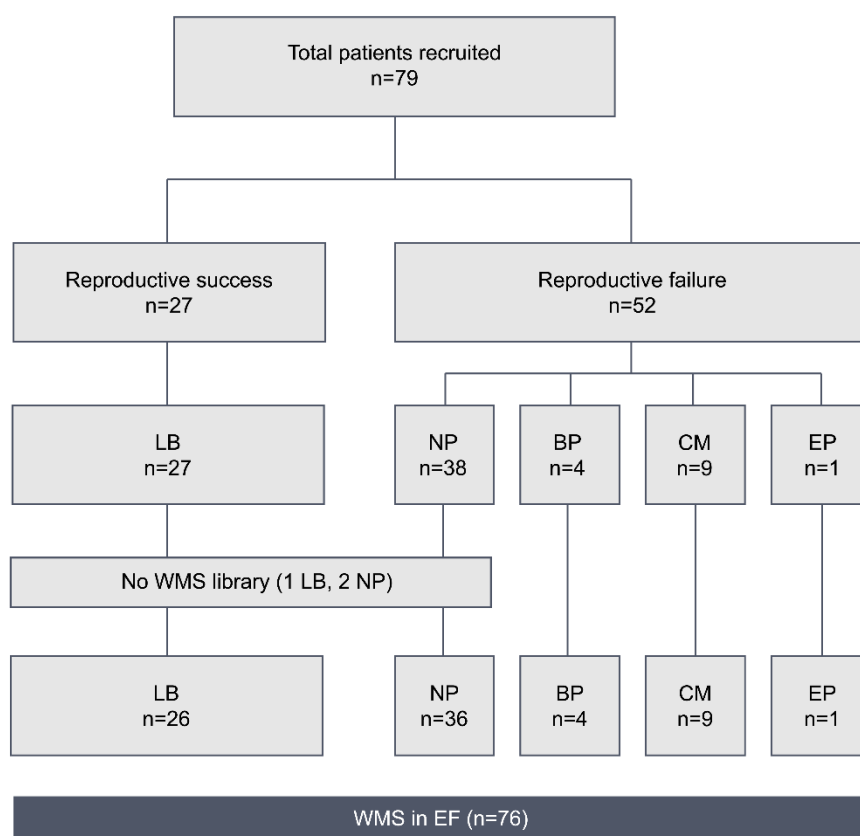


Figure 26. Patients allocated in each reproductive outcome group after library preparation. Abbreviations: **BP**, biochemical pregnancy; **CM**, clinical miscarriage; **EB**, endometrial biopsy; **EF**, endometrial fluid; **EP**, ectopic pregnancy; **LB**, live birth; **NP**, no pregnancy; **WMS**, whole metagenomic sequencing.

The metagenomic analysis of EF samples using WMS yielded 5,633,767,812 paired reads. Applying quality control parameters provided a mean of 38,091,050 joined sequencing reads per sample (range 3,453,408 - 595,255,010) with an average size of 216 bp. After filtering out human reads, we obtained a mean of 303,679 sequencing reads per sample (range 18,707 – 12,074,722); this data demonstrated that only ~0.5-1% of reads correspond to microbial DNA, in agreement with previous reports (Moreno et al., 2020).

Based on the composition and abundance of the taxa found in each sample, bacterial communities were then classified into different community state types (CSTs)

as previously described (Ravel et al., 2011). We encountered seven CSTs in EF samples, while only four of them were found in more than one sample (Figure 27).

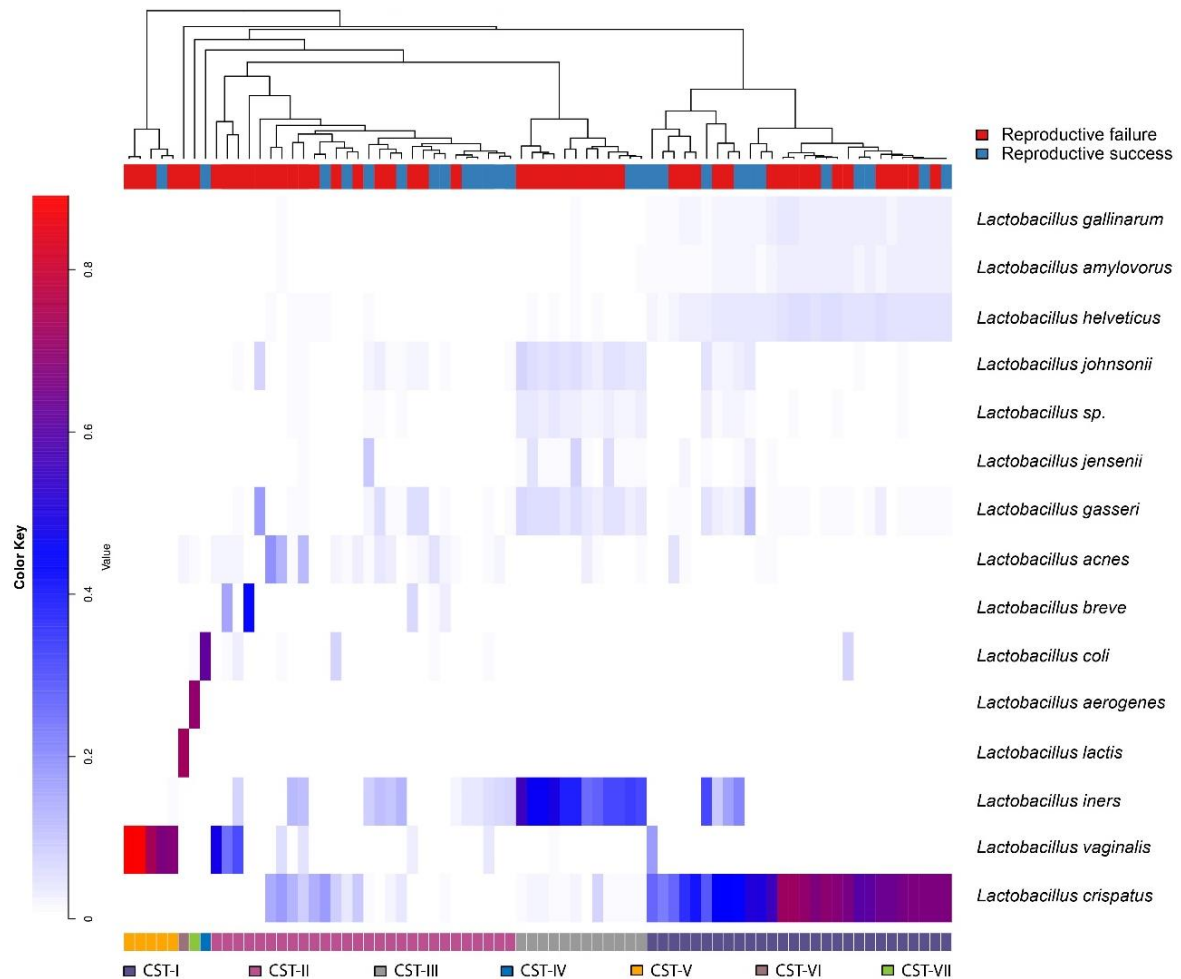


Figure 27. Distribution of CSTs in endometrial samples from patients with distinct reproductive outcomes. Four major and three minor community state types (CST-I, CST-II, CST-III, and CST-V) described in endometrial samples of women undergoing fertility treatment.

CST-I was present in 28 samples and displayed high proportions of *L. crispatus* and minor presence of commensals mainly from the *Lactobacillus* genus, as well as low numbers of dysbiotic bacteria (Figure 28A). CST-II comprised 28 samples with a highly homogeneous presence of several commensal bacteria including *Lactobacillus* spp.

(*L. iners*, *L. gasseri*, *L. johnsonii*, *L. crispatus*, among others), *Cutibacterium acnes*, and minor presence of other commensals and pathogenic bacteria such as *E. coli* and *G. vaginalis* (Figure 28B). CST-III, comprising 12 samples, was predominantly composed of *L. iners*, with minor abundance of several other *Lactobacilli*, commensals, and pathogens (Figure 28C). Finally, 5 endometrial samples belonged to CST-V, mainly composed of *G. vaginalis*, one of the most common pathogens in the reproductive tract, alongside scarce percentages of *Bifidobacterium breve*, *L. iners*, and others (Figure 28D). Additionally, 3 other CSTs were identified with only one sample in each. The predominant bacteria in these CSTs were *E. coli*, *Lactococcus lactis*, and *Klebsiella aerogenes* in CST-IV, -VI, and -VII, respectively (Figure 28E-G).

We next aimed to assess whether endometrial microbiota composition associates with reproductive outcomes, looking for an association between CST composition and reproductive outcomes in terms of reproductive success (LB) or failure (NP, BP, CM, and EP combined) (Figure 29).

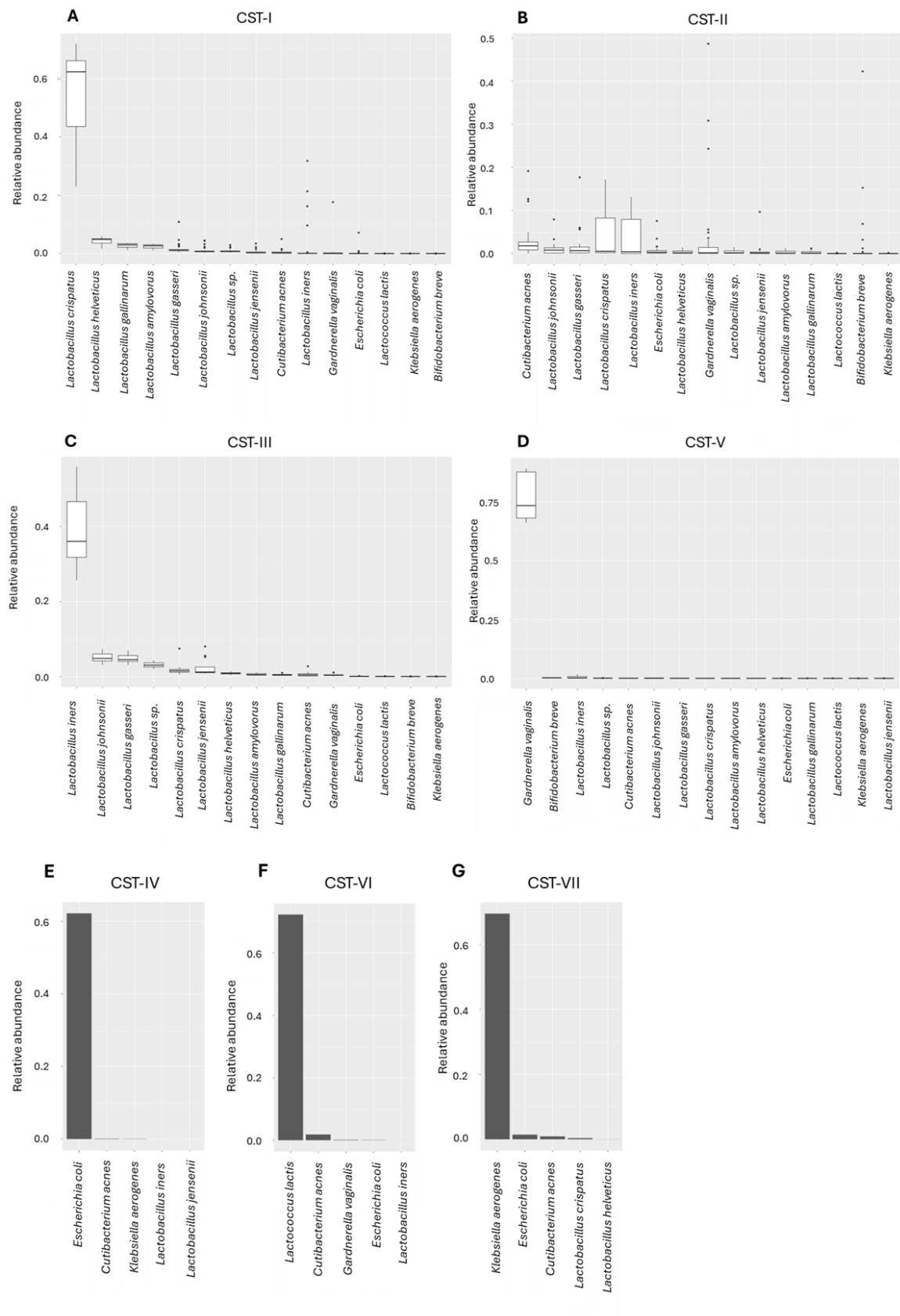


Figure 28. Taxonomic composition of the samples in each community state type. Top fifteen (A-D) and top five (E-G) most common bacterial species that compose each of the described community state types.

Importantly, patients whose bacterial communities were dominated by *L. iners* (CST-III) or *G. vaginalis* (CST-V) mostly had failed cycles (83.3% and 80.0%, respectively) (Figure 29). On the other hand, samples dominated by *L. crispatus* (CST-I) and those that displayed a heterogeneous mix of *Lactobacilli* (CST-II) generally had better reproductive outcomes (LB rates, both 39.3%) compared to CST-III and -V (LB rates, 16.7% and 20.0%) (Figure 29).

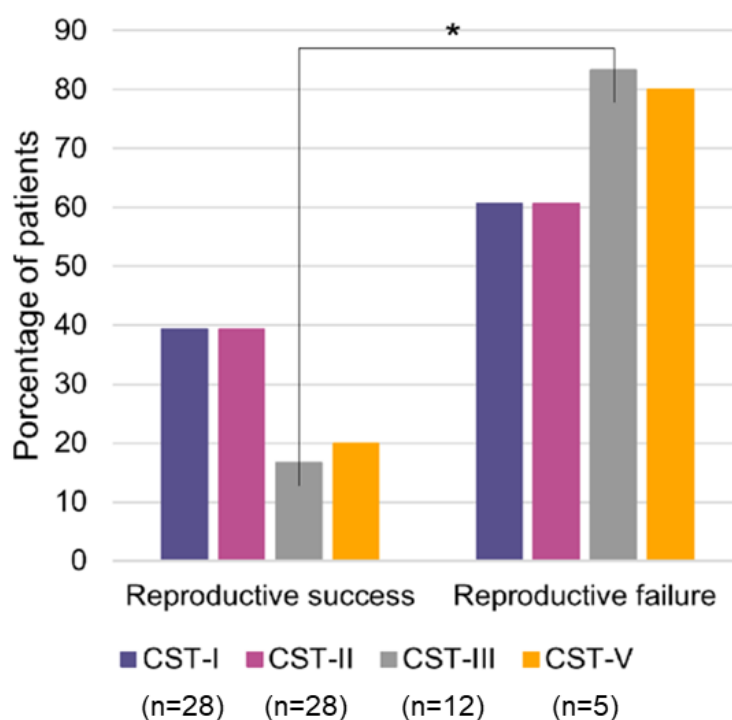


Figure 29. Differential presence of CSTs between reproductive success and failure. Bar plot depicting the percentage reproductive outcomes in each CST. Individual proportion test, *indicates p-value <0.05.

Still, when we compared the incidence of CSTs between reproductive outcomes, no significant differences were observed (Fisher's exact test, $p > 0.05$). When checking for differences in the proportions of reproductive outcomes within CST-I, -II and -III, only CST-III displayed significant results (Individual proportion test, $p = 0.043$), while CSTs-I, and -II had no statistically significant proportion differences (Individual proportion test, $p > 0.05$) (Figure 29). Due to the small sample size in CST-IV, -V, -VI,

and -VII (n=1, 5, 1 and 1, respectively), there was not enough statistical power to assess differences within each of these CSTs.

Nonetheless, it is important to highlight that cases of reproductive failure were found distributed across all bacterial communities (Figures 27 and 29). Additionally, of the 8 patients that exhibited a *G. vaginalis* abundance exceeding 20%, 7 had a failed reproductive cycle (5 NP, 1 BP, and 1 CM), while one of them – presenting a CST-V microbiome – ended up being part of the 26 women that reached a reproductive success (LB) (Figure 27 and Table 5). Thus, although these data support previous evidence that endometrial bacterial abundance associates with reproductive outcomes, overall microbiota composition not always associates with poor or favorable reproductive outcomes; suggesting strain interaction and host factors might come into play and modulate microbiome impact on endometrial physiology at the molecular or cellular level.

Table 5. *G. vaginalis* abundance in endometrial samples.

Sample	<i>G. vaginalis</i> abundance (%)	Reproductive outcome	CST
A	1.070	NP	3
B	3.808	LB	2
C	3.233	NP	2
D	5.579	NP	2
E	71.598	LB	5
F	77.479	NP	5
G	49.122	CM	2
H	31.313	BP	2
I	70.902	NP	5
J	24.535	NP	2
K	89.922	NP	5
L	94.584	NP	5
M	5.463	NP	2
N	19.782	LB	1

Only samples with *G. vaginalis* abundance exceeding 1% are included in the table. Bold type highlights abundance values above 20%. Abbreviations: **BP**, biochemical pregnancy; **CM**, clinical miscarriage; **EP**, ectopic pregnancy; **LB**, live birth; **NP**, no pregnancy.

5.1.1 Bacterial species associated with reproductive outcomes

When assessing the abundance of individual bacterial species, findings agree with those observed in previous studies using 16S rRNA analysis (Moreno et al., 2022). WMS results pointed to an increased abundance of dysbiotic and pathogenic taxa associated with reproductive failure; a feature that was now observed at the species level. Specifically, bacterial pathogens causing BV (e.g., *G. vaginalis* and *A. vaginae*) and CE (e.g., *Streptococcus agalactiae*, *Enterococcus faecalis*, and *Staphylococcus aureus*) displayed a significant association with reproductive failure, as well as *Fusobacterium nucleatum*, several species of *Bifidobacterium*, and others (Table 6).

Table 6. Taxa significantly enriched in reproductive failure compared to reproductive success.

Taxa	Log2 Fold Change	p-adj
<i>Gardnerella</i>	5.45	0.00
<i>Gardnerella</i> sp.	5.25	0.00
<i>Bifidobacterium breve</i>	4.90	0.00
<i>Gardnerella vaginalis</i>	4.73	0.00
<i>Bifidobacterium longum</i>	4.68	0.00
<i>Alloscardovia omnicolens</i>	3.34	0.00
<i>Olsenella</i> sp.	3.27	0.00
<i>Bifidobacterium bifidum</i>	3.26	0.00
<i>Lactobacillus parafarraginis</i>	3.21	0.00
<i>Atopobium vaginae</i>	3.12	0.00
<i>Lactobacillus coleohominis</i>	2.93	0.00
<i>Bifidobacterium</i>	2.89	0.00
<i>Lactobacillus rennini</i>	2.71	0.00
<i>Lactobacillus plantarum</i>	2.71	0.00
<i>Lactobacillus malefermentans</i>	2.63	0.00
<i>Lactobacillus brevis</i>	2.63	0.00
<i>Streptococcus</i>	2.60	0.00
<i>Lactobacillus collinoides</i>	2.48	0.00
<i>Lactobacillus sunkii</i>	2.46	0.00
<i>Staphylococcus epidermidis</i>	2.45	0.00
<i>Alloscardovia macacae</i>	2.42	0.00
<i>Streptococcus</i> sp.	2.39	0.00
<i>Lactobacillus gastricus</i>	2.38	0.00
<i>Veillonellaceae bacterium</i>	2.37	0.00
<i>Lactobacillus hominis</i>	2.33	0.02

Taxa	Log2 Fold Change	p-adj
<i>Lactobacillus oris</i>	2.31	0.00
<i>Lactococcus lactis</i>	2.31	0.00
<i>Streptococcus dysgalactiae</i>	2.31	0.00
<i>Staphylococcus aureus</i>	2.30	0.00
<i>Alloscardovia</i>	2.28	0.00
<i>Scardovia wiggsiae</i>	2.27	0.00
<i>Lactobacillus reuteri</i>	2.26	0.01
<i>Enterococcus faecalis</i>	2.26	0.00
<i>Haemophilus parainfluenzae</i>	2.26	0.00
<i>Lactobacillus taiwanensis</i>	2.25	0.02
<i>Lactobacillus otakiensis</i>	2.25	0.00
<i>Fusobacterium nucleatum</i>	2.23	0.00
<i>Streptococcus salivarius</i>	2.20	0.00
<i>Lactobacillus psittaci</i>	2.17	0.03
<i>Lactobacillus parabuchneri</i>	2.17	0.01
<i>Corynebacterium kroppenstedtii</i>	2.17	0.00
<i>Staphylococcus hominis</i>	2.14	0.00
<i>Streptococcus cristatus</i>	2.12	0.00
<i>Lactobacillus equicursoris</i>	2.11	0.03
<i>Lactobacillus paralimentarius</i>	2.11	0.01
<i>Streptococcus equinus</i>	2.09	0.00
<i>Lactobacillus panis</i>	2.05	0.02
<i>Mobiluncus mulieris</i>	2.04	0.00
<i>Streptococcus intermedius</i>	2.03	0.00
<i>Coriobacterium glomerans</i>	2.01	0.00
<i>Lactobacillus kalixensis</i>	2.00	0.04
<i>Lactobacillus sakei</i>	2.00	0.01
<i>Streptococcus agalactiae</i>	2.00	0.01

Taxonomical profiles were obtained using Diamond/BLAST pipelines and processed using R statistical software. Significant enrichment was considered when Log2 Fold Change was >2 and adjusted p-value (p-adj) <0.05.

5.1.2 Taxonomic networks and microbe coexistence

Next, we aimed to examine the impact of the overall microbiota community viewing bacterial species as integral components of a complex system. By adopting an ecological perspective, we sought to further characterize bacterial interactions within the uterine microbiota and its association to reproductive outcomes. Thus, we next constructed taxonomic networks to assess whether the abundance of specific species associates with presence/absence of other bacteria.

Taxonomic co-occurring bacterial plots displayed significant differences regarding reproductive outcomes: weak networks were observed in LB and NP, whilst BP and CM displayed more complex associations (Figures 30-37).

Among samples from the LB group, we observed positive associations of *L. crispatus* with several species of *Lactobacillus* but not *L. iners*, while *L. iners* only co-existed with *L. hominis*. Additionally, a small, isolated node comprising positive associations between several *Bifidobacterium* species was also found (Figures 30 and 31).

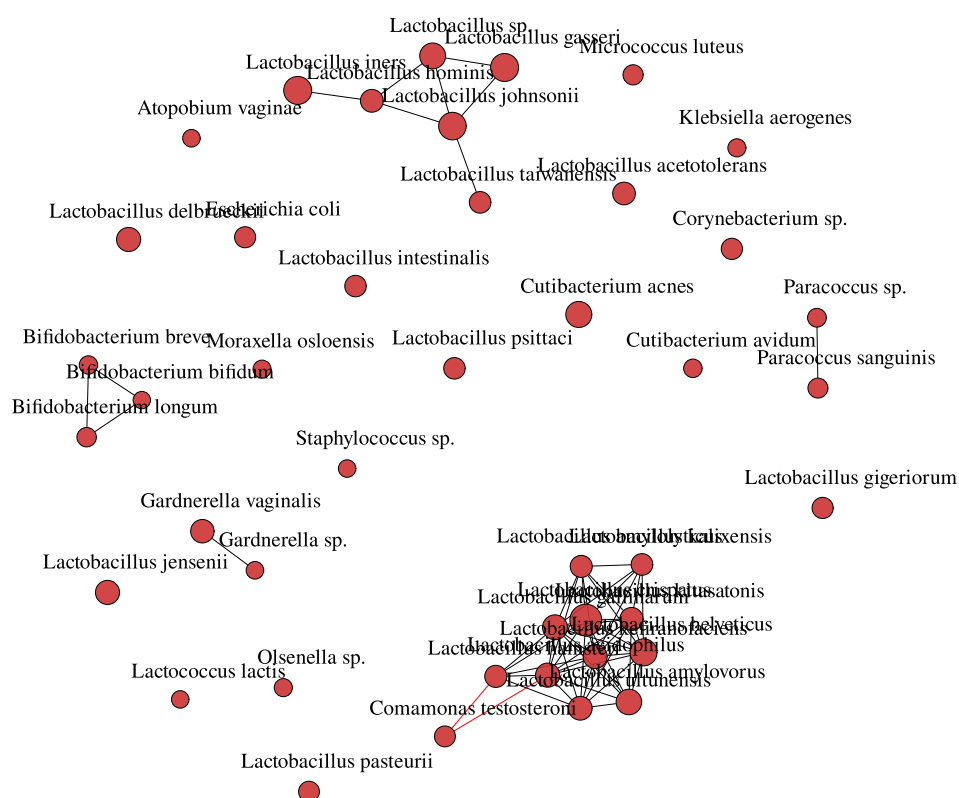


Figure 30. Taxonomic networks within samples from the LB group. Network analysis reveals microbe co-occurrence in endometrial samples. Node properties: circle size, proportional to the normalized and standardized relative abundances. Edge properties: red for negative and grey for positive correlation coefficients (values >0.7 and <-0.7 ; p -value <0.05).

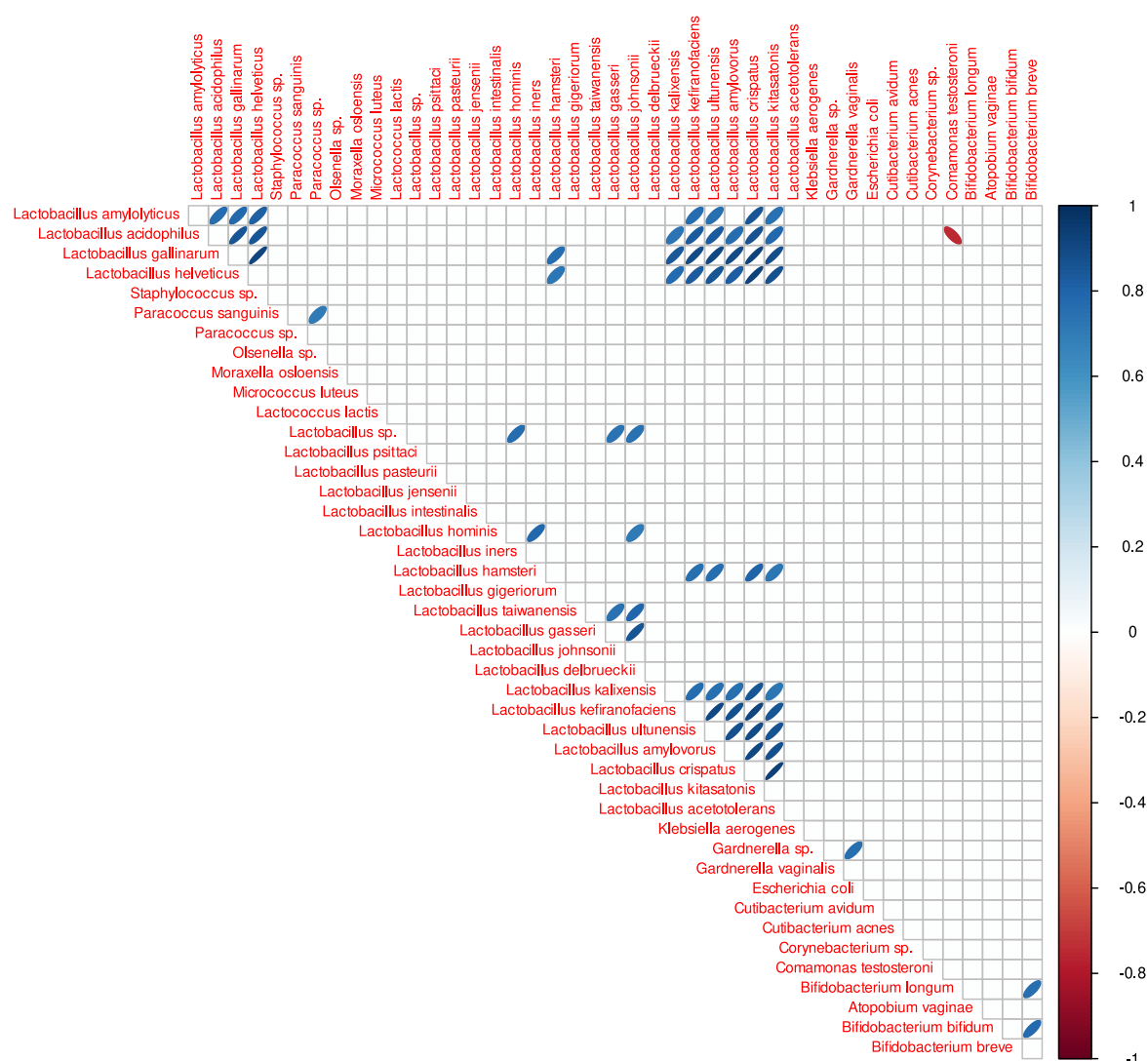


Figure 31. Microbe coexistence within samples from the LB group. SparCC taxonomic correlation plots showed significant interactions. Correlation plot depicting coexisting bacteria within endometrial samples. Edge properties: red for negative correlation coefficients, blue for positive correlation coefficients; thickness is proportional to correlation significance.

In NP, two major nodes appeared: the first one, in which several *Lactobacilli* including *L. crispatus*, *L. acidophilus*, and *L. helveticus* displayed positive associations among them, and a second node enriched with positive associations between *G. vaginalis*, *Bifidobacterium* spp., *A. vaginae*, and *Olsenella* sp. Additionally, a smaller node comprised *Cutibacterium* spp., *Micrococcus luteus*, and *Comamonas testosteroni*, species not commonly reported in the uterine cavity. Interestingly, *L. iners*

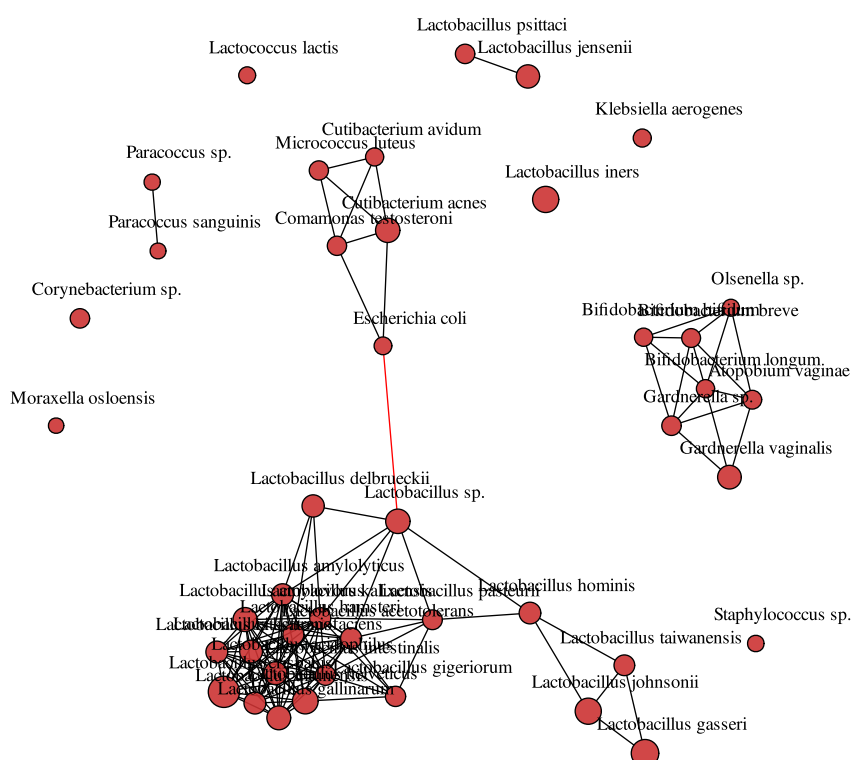


Figure 32. Taxonomic networks within samples from the NP group. Network analysis reveals microbe co-occurrence in endometrial samples. Node properties: circle size, proportional to the normalized and standardized relative abundances. Edge properties: red for negative and grey for positive correlation coefficients (values >0.7 and <-0.7 ; p -value <0.05).

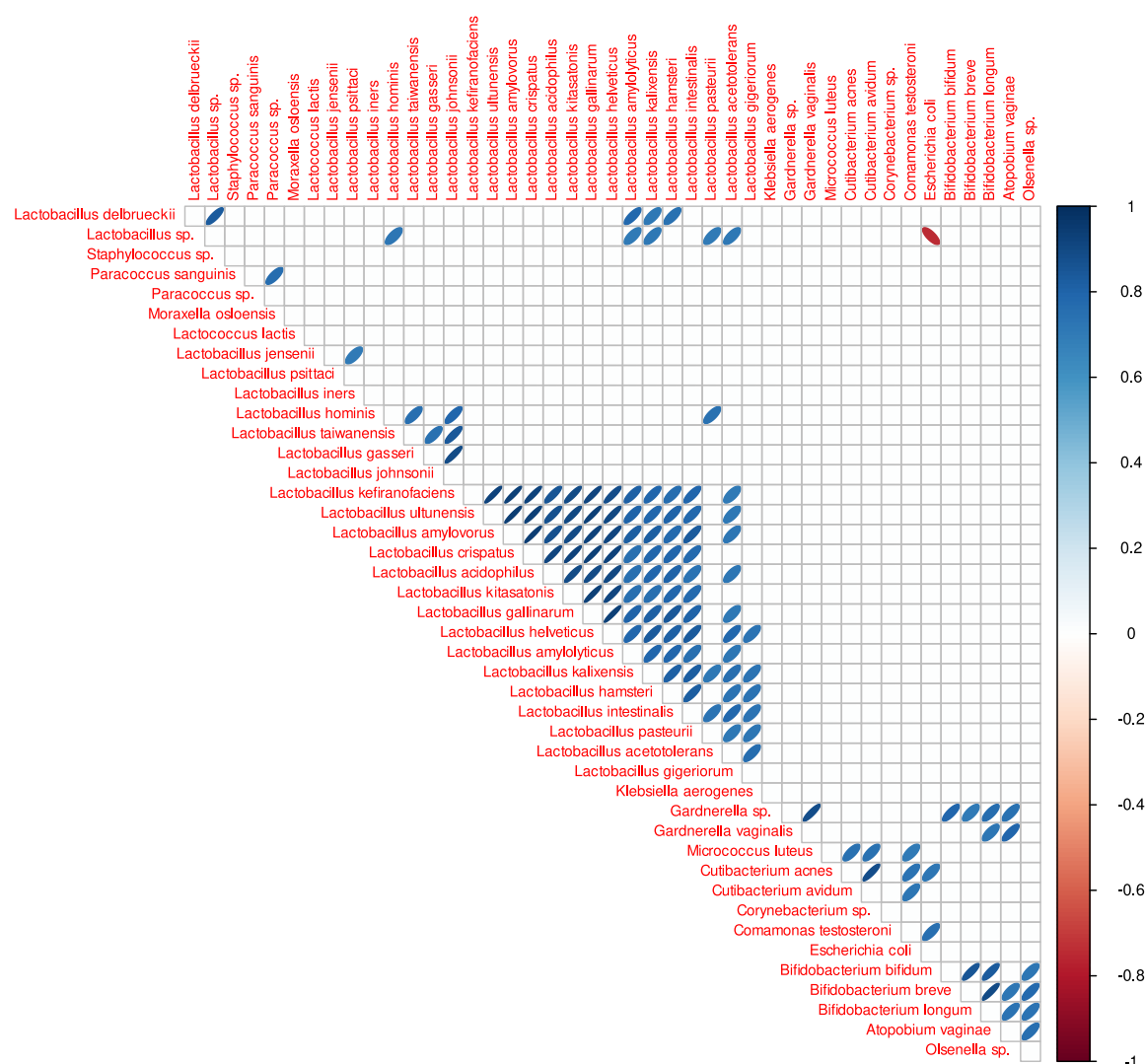


Figure 33. Microbe coexistence within samples from the NP group. Correlation plot depicting coexisting bacteria within endometrial samples. Edge properties: red for negative correlation coefficients, blue for positive correlation coefficients; thickness is proportional to correlation significance.

In BP and CM, more complex networks were found, suggesting a more disruptive environment in these groups (Figures 34-37).

Among several associations in the BP group, we highlight a big block of *Lactobacilli* including *L. crispatus*, *L. acidophilus*, *L. helveticus*, and *L. delbrueckii* displaying positive associations between them and negative associations with pathogens and dysbiotics such as *K. aerogenes*, *A. vaginae*, *Bifidobacterium* spp., and *G. vaginalis*. Another node included positive associations between *Paracoccus* spp., *Micrococcus*

luteus, and *Moraxella osloensis*, which correlated negatively with *E. coli*, *Bifidobacterium* spp., and *Gardnerella*. Finally, exclusively in this group, *L. iners* was found to coexist with *Cutibacterium* spp., *A. vaginae*, *K. aerogenes*, *G. vaginalis*, and a few other *Lactobacilli*; but also associated negatively with a major group of *Lactobacillus* comprising *L. crispatus* and others. (Figures 34 and 35).

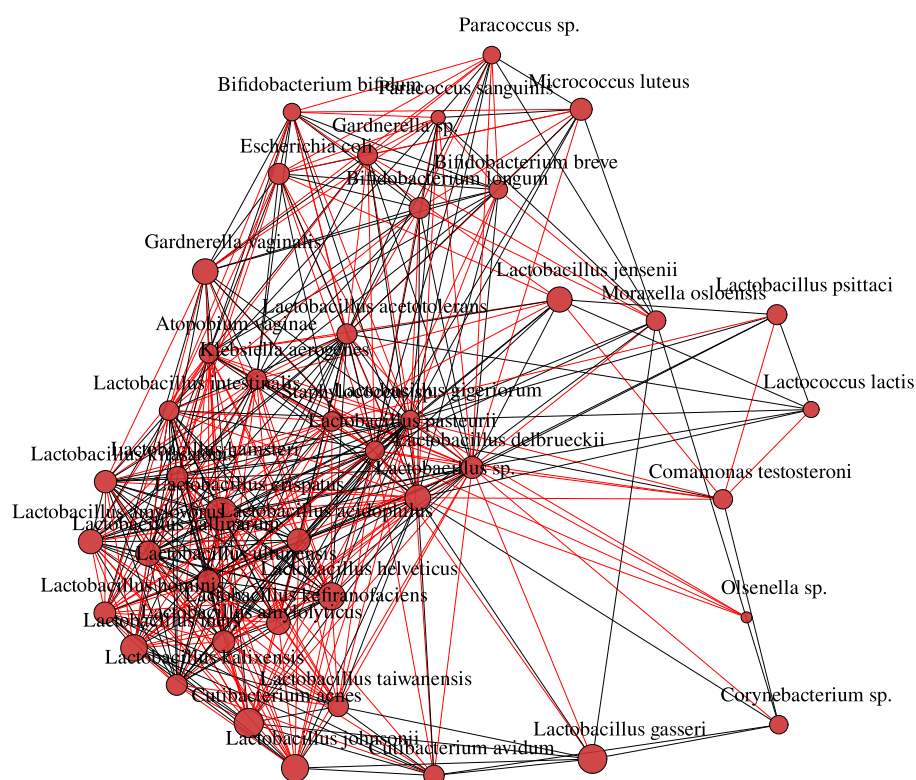


Figure 34. Taxonomic networks within samples from the BP group. Network analysis reveals microbe co-occurrence in endometrial samples. Node properties: circle size, proportional to the normalized and standardized relative abundances. Edge properties: red for negative and grey for positive correlation coefficients (values >0.7 and <-0.7 ; p -value <0.05).

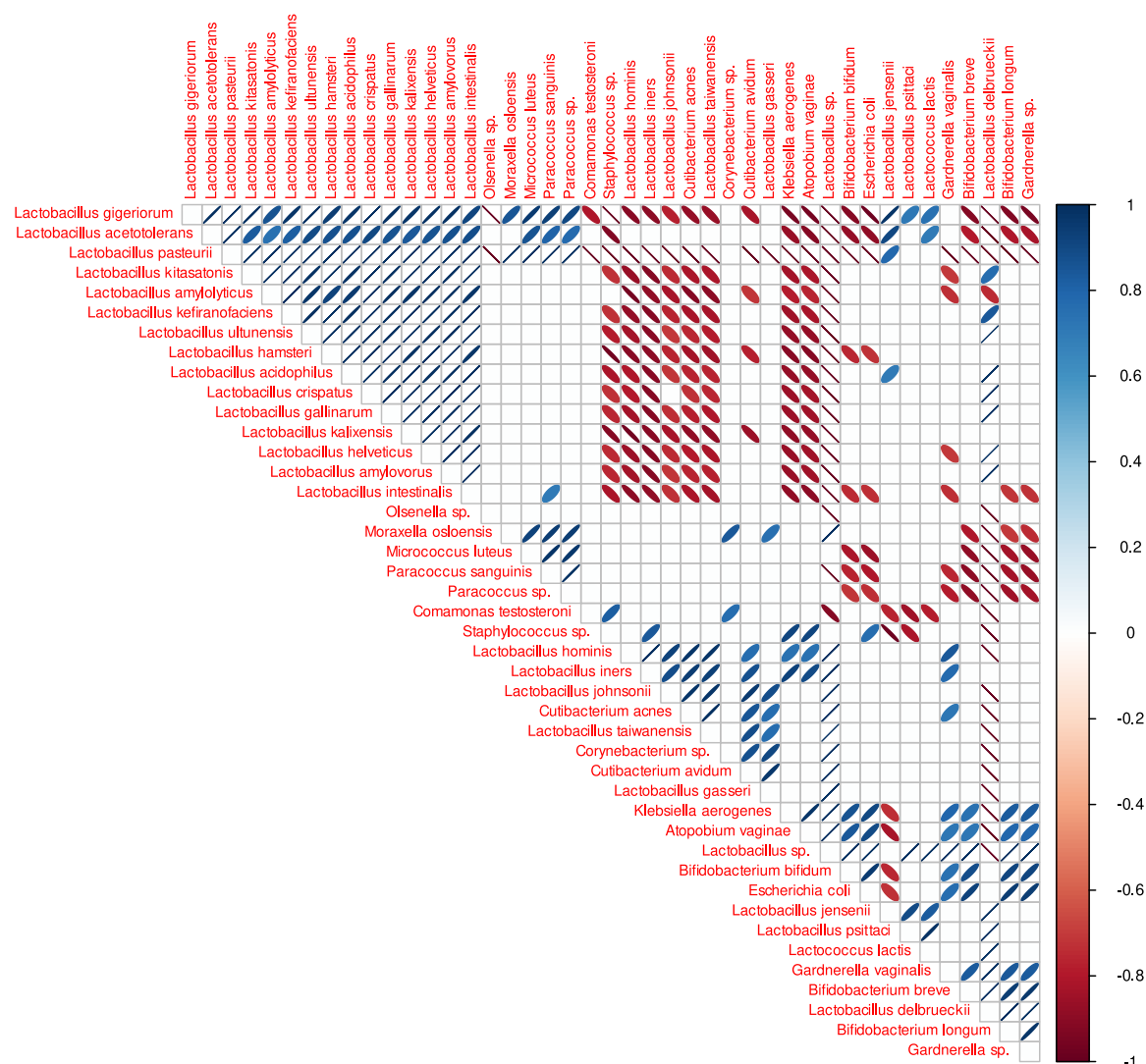


Figure 35. Microbe coexistence within samples from the BP group. Correlation plot depicting coexisting bacteria within endometrial samples. Edge properties: red for negative correlation coefficients, blue for positive correlation coefficients; thickness is proportional to correlation significance.

Lastly, the CM group also displayed a complex taxonomic network in which multiple *Lactobacilli* coexist but associated negatively with *A. vaginae* and *G. vaginalis*, which coexisted with *Olsenella* sp., *Bifidobacterium* spp., and *Lactococcus lactis*. A third node included positive associations between *Micrococcus luteus*, *Cutibacterium* spp., *Moraxella osloensis*, and *Paracoccus* spp., which associated negatively with some *Lactobacillus* species such as *L. jensenii*, *L. johnsonii*, and *L. gasseri*. In this group, *L. iners* had an opposite behavior and associated positively with few other *Lactobacilli*,

including *L. gasseri* and *L. johnsonii*, and did not associate directly with pathogens nor dysbiotic bacteria (Figures 36 and 37).

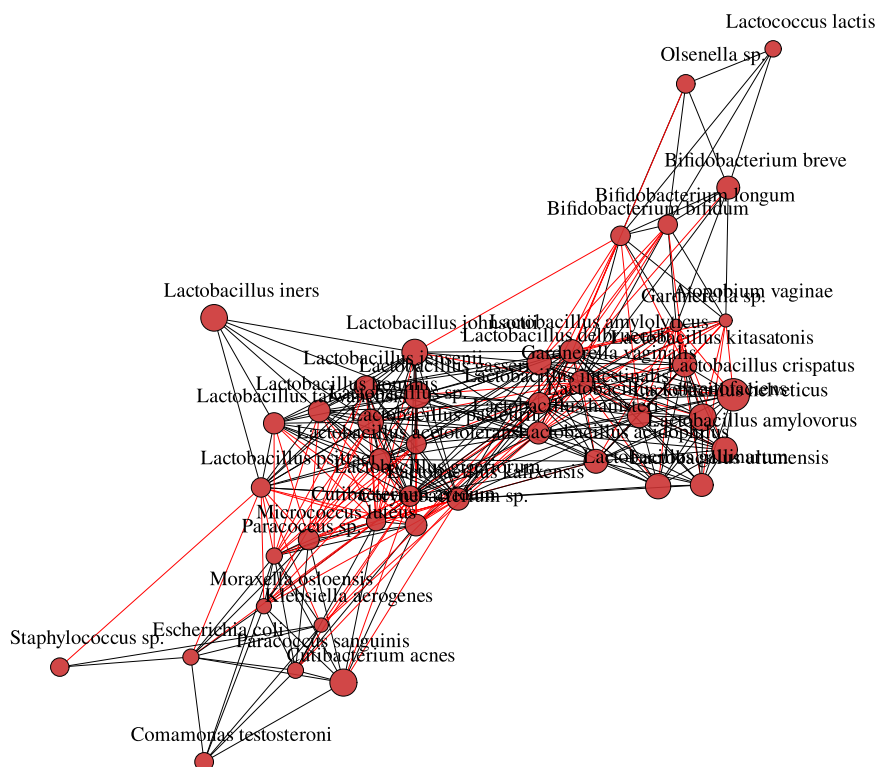


Figure 36. Taxonomic networks within samples from the CM group. Network analysis reveals microbe co-occurrence in endometrial samples. Node properties: circle size, proportional to the normalized and standardized relative abundances. Edge properties: red for negative and grey for positive correlation coefficients (values >0.7 and <-0.7 ; p -value <0.05).

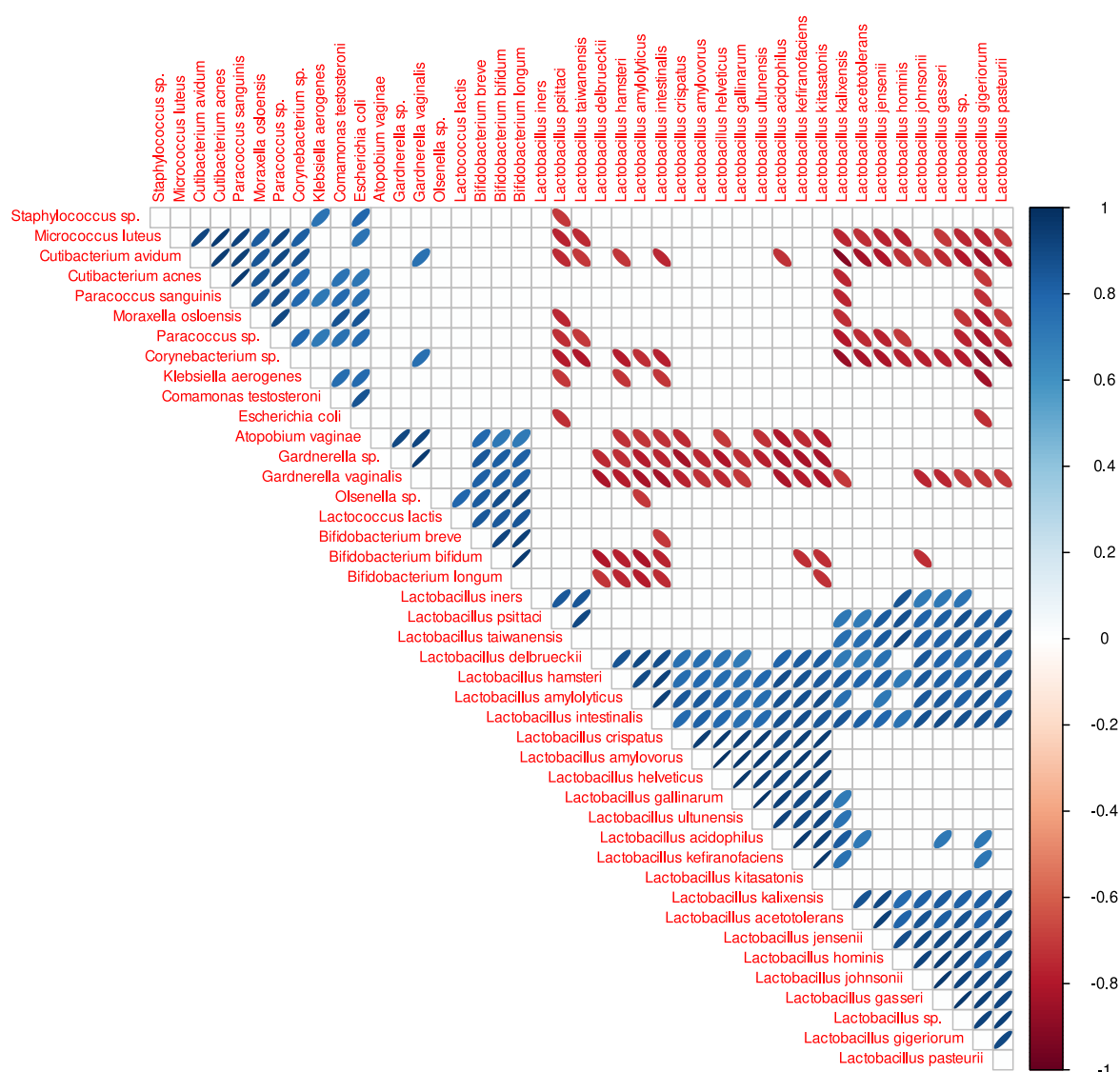


Figure 37. Microbe coexistence within samples from the CM group. Correlation plot depicting coexisting bacteria within endometrial samples. Edge properties: red for negative correlation coefficients, blue for positive correlation coefficients; thickness is proportional to correlation significance.

A significant difference between the networks observed for the different reproductive outcomes involves *G. vaginalis*, which was an isolated taxon in LB, but presented differential positive and negative associations with other taxa in reproductive failure. For instance, it positively associated with *Bifidobacterium* spp. in NP, showed positive associations with *L. iners*, *A. vaginae*, *K. aerogenes*, *B. bifidum*,

and *E. coli* in BP, while presented negative association with *L. crispatus* and concomitant positive associations with *Bifidobacterium* spp., and *A. vaginae* in CM.

Overall, these results are the first to describe the distinct CSTs found within the endometrial microbiota; and demonstrate that bacterial communities dominated by *L. iners* or *G. vaginalis* tend to associate with poorer reproductive outcomes. Additionally, we corroborate previous findings that support an association between the presence of dysbiotic and/or pathogenic genera and poor reproductive outcomes, as we identified, at the species level, which specific bacteria are mostly associated with reproductive failure. By building networks and maps of coexistence, these data might help understand the impact of whole bacterial communities in reproductive outcomes.

5.2 GENETIC AND FUNCTIONAL DIVERSITY OF BACTERIAL COMMUNITIES IN REPRODUCTIVE OUTCOMES

After deeply assessing the relevance of endometrial bacterial communities at the taxonomic level, we next sought to characterize how different bacterial communities contributed to the uterine environment at the molecular level. Thus, genetic and functional diversity of bacterial communities in different reproductive outcomes were assessed through WMS and Diamond/BLAST pipelines.

5.2.1 Bacterial physiological functions

To further investigate the impact the microbiome may have upon endometrial cells, we next aimed to assess which bacterial genes and functions were enriched in each reproductive outcome by analyzing global metagenomics data. Using DESeq2, we identified fourteen genes significantly decreased in the NP compared to the LB group. These genes are associated with several catalytic reactions and biological processes

such as detoxication (glutathione S-transferase family protein, $p\text{-adj}=0.03$), amino acids processing (methionine synthase, $p\text{-adj}=0.01$), iron ion binding (aromatic ring-hydroxylating dioxygenase subunit alpha, $p\text{-adj}=0.02$), protein transport (nuclear transport factor 2 family protein, $p\text{-adj}=0.03$), among others (Table 7).

Table 7. Bacterial genes enriched in the live birth group compared to the no pregnancy group using DeSeq2 analysis.

Gene	Base Mean	log2 FC	p-adj
MULTISPECIES: sulfite oxidase-like oxidoreductase	7,551.00	4.77	0.02
Aromatic ring-hydroxylating dioxygenase subunit alpha	12.55	4.45	0.02
Glutathione S-transferase family protein	6,778.00	4.44	0.03
MULTISPECIES: Holliday junction branch migration DNA helicase RuvB	10,936.00	3.54	0.00
4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	11.48	3.37	0.03
MULTISPECIES: glycine cleavage system protein T	4.66	3.31	0.03
MULTISPECIES: cation acetate symporter	7.07	3.22	0.02
Sulfurtransferase FdhD	6,181.00	3.2	0.03
Acetate-CoA ligase	7,353.00	3.11	0.03
Methionine synthase	24,327.00	2.92	0.01
MoxR family ATPase	7,497.00	2.83	0.03
Nuclear transport factor 2 family protein	13,783.00	2.66	0.03
MULTISPECIES: type I methionyl aminopeptidase	12,998.00	2.39	0.02
MULTISPECIES: elongation factor 4	12,175.00	2.19	0.02

Abbreviations: **FC**, fold change; **p-adj**, adjusted p value.

Considering DESeq2 is a strict statistical tool which performance relies on a high number of samples or replicates, our low number of samples in BP and CM groups did not allow for proper analysis. Thus, we next used the LefSe approach to identify significantly enriched microbial genes in each reproductive outcome (Table 8). Compared to NP, BP, and CM, discriminatory genes in the LB group mainly included those involved in metabolism pathways such as oxidoreductase and adenine phosphoribosyltransferase ($p=0.04$ and $p=0.01$), but also DNA polymerase ($p=0.04$) and others. Samples from the NP group displayed enriched presence of DNA

polymerase III subunit beta, preprotein translocase subunit SecY and DUF340 domain-containing protein ($p=0.03$, $p=0.04$, and $p=0.04$, respectively), related to genetic processing and environmental adaptation. Samples from the BP group mainly displayed enrichment of genes related to metabolism such as dihydrodipicolinate synthase ($p=0.01$), propionyl-CoA synthetase ($p=0.01$), and gamma-glutamyl phosphate reductase ($p=0.02$), among other genes. Lastly, the CM group displayed the highest number of enriched genes, also comprising a wider range of functions such as environmental information processing (sugar ABC transporter substrate-binding protein, $p=0.02$), metabolism (threonine-tRNA ligase, $p=0.00$), and genetic information processing (50S ribosomal protein L11, $p=0.03$) (Table 8).

Table 8. LEfSe analysis describing bacterial genes enriched in each reproductive outcome.

Enriched in	Gene	Logarithm value	p-value
LB	Oxidoreductase	3.35	0.04
	Adenine phosphoribosyltransferase	3.10	0.01
	Sulfurtransferase FdhD	2.63	0.03
	DNA polymerase	2.52	0.04
	WxL domain-containing protein	2.45	0.01
	MoxR family ATPase	2.45	0.02
	1-aminocyclopropane- 1-carboxylate deaminase	2.33	0.02
NP	DNA polymerase III subunit beta	3.09	0.03
	Preprotein translocase subunit SecY	3.03	0.04
	DUF340 domain-containing protein	2.53	0.04
BP	Dihydrodipicolinate synthase family protein	2.54	0.01
	DUF2815 domain-containing protein	2.48	0.02
	Propionyl-CoA synthetase	2.47	0.01
	ATP dependent helicase HrpB	2.46	0.03
	NADPH-dependent 2,4-dienoyl-CoA reductase	2.41	0.02
	Gamma-glutamyl phosphate reductase	2.41	0.02
	Polyketide cyclase	2.39	0.02
	MULTISPECIES: Acetyl-CoA acetyltransferase	2.39	0.03
	NADP dependent oxidoreductase	2.38	0.01
	Glutamate-5-semialdehyde dehydrogenase	2.32	0.01
	DUF1768 domain-containing protein	2.18	0.04

Enriched in	Gene	Logarithm value	p-value
CM	Sugar ABC transporter substrate-binding protein	3.35	0.02
	Threonine-tRNA ligase	3.33	0.00
	Bifunctional folypolyglutamate synthase-dihydrofolate synthase	3.00	0.03
	AAA family ATPase	2.98	0.04
	GTPase Era	2.73	0.01
	50S ribosomal protein L11	2.72	0.03
	Cystathionine gamma-synthase	2.71	0.02
	30S ribosomal protein S6	2.67	0.01
	Transcriptional regulator NrdR	2.63	0.03
	Peptidase E	2.55	0.00
	MULTISPECIES: ferrous iron transporter B	2.43	0.00
	MULTISPECIES: 50S ribosomal protein L11	2.42	0.04
	Isocitrate dehydrogenase (NADP)	2.39	0.01
	MULTISPECIES: transferase	2.37	0.03
	Peptidase S1	2.37	0.01
	Fumarate reductase-quinol flavoprotein subunit	2.35	0.00
	Alkyl hydroperoxide reductase	2.29	0.04
	MULTISPECIES: ATPase AAA	2.29	0.04
	Glycogen/starch/Alpha-glucan family phosphorylase	2.28	0.04
	DUF1541 domain-containing protein	2.25	0.00
	Acyl_CoA thioesterase II	2.25	0.02
	Type VI secretion protein VskK	2.33	0.04
	1,3-beta-galactosyl-N-acetylhexosamine phosphorylase	2.19	0.02

Abbreviations: **BP**: biochemical pregnancy; **CM**: clinical miscarriage; **NP**: no pregnancy; **LB**: live birth.

Taking results of both methods into consideration, we observed different abundance of several genes between groups (Tables 9-12). In comparison to NP, BP, and CM, LB samples displayed enrichment of genes related to the production of amino acids and nucleotides, the metabolism of vitamins and cofactors, xenobiotics metabolism, DNA replication and repair, and translation (Table 9).

Table 9. Bacterial functions and associated genes significantly enriched in LB.

Class	Function	Pathway	Gene
Genetic information processing	DNA replication	DNA polymerase	DNA polymerase
	Replication and repair	DNA helicase	<i>MULTISPECIES: Holliday junction branch migration DNA helicase RuvB</i>
		Homologous recombination	<i>MULTISPECIES: Holliday junction branch migration DNA helicase RuvB</i>
	Translation	Elongation	<i>MULTISPECIES: elongation factor 4</i>
		Translation, ribosomal structure, and biogenesis	<i>MULTISPECIES: type I methionyl aminopeptidase</i>
Metabolism	Amino acid metabolism	Glycine, serine, and threonine metabolism	<i>MULTISPECIES: glycine cleavage system protein T</i>
			<i>Methionine synthase</i>
		Cysteine and methionine metabolism	<i>Methionine synthase</i>
			<i>1-aminocyclopropane- 1-carboxylate deaminase</i>
		Tyrosine metabolism	<i>Glutathione S-transferase family protein</i>
	Carbohydrate metabolism	Glycolysis /Gluconeogenesis	<i>Acetate-CoA ligase</i>
		Glyoxylate and dicarboxylate metabolism	<i>Acetate-CoA ligase</i>
		Propanoate metabolism	<i>Acetate-CoA ligase</i>
		Pyruvate metabolism	<i>Acetate-CoA ligase</i>
		Hydrocarbon metabolism	<i>Aromatic ring-hydroxylating dioxygenase subunit alpha</i>
		Glyoxylate and dicarboxylate metabolism	<i>MULTISPECIES: glycine cleavage system protein T</i>
	Energy metabolism	Carbon fixation pathways in prokaryotes	<i>Acetate-CoA ligase</i>
			<i>Methionine synthase</i>
		Methane metabolism	<i>Acetate-CoA ligase</i>
			<i>Methionine synthase</i>
			<i>MULTISPECIES: glycine cleavage system protein T</i>
		Sulfur metabolism	Sulfurtransferase FdhD
	Global metabolism	Biosynthesis of secondary metabolites	<i>4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase</i>
			<i>Acetate-CoA ligase</i>
			<i>Methionine synthase</i>
		Biosynthesis of amino acids	<i>Methionine synthase</i>
			<i>Methionine synthase</i>
		Carbon metabolism	<i>MULTISPECIES: glycine cleavage system protein T</i>
			<i>Acetate-CoA ligase</i>
			<i>Acetate-CoA ligase</i>
		Microbial metabolism in diverse environments	<i>Methionine synthase</i>
		Other	<i>Oxidoreductase</i>
		Starch and sucrose metabolism	<i>Acetate-CoA ligase</i>

Class	Function	Pathway	Gene
Metabolism	Metabolism of cofactors and vitamins	One carbon pool by folate	<i>MULTISPECIES: glycine cleavage system protein T</i>
			<i>Methionine synthase</i>
		Porphyrin metabolism	MoxR family ATPase
	Metabolism of other amino acids	Glutathione metabolism	<i>Glutathione S-transferase family protein</i>
		Selenocompound metabolism	<i>Methionine synthase</i>
	Metabolism of terpenoids and polyketides	Terpenoid backbone biosynthesis	<i>4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase</i>
	Nucleotide metabolism	Purine metabolism	Adenine phosphoribosyltransferase
	Xenobiotics biodegradation and metabolism	Chlorocyclohexane and chlorobenzene degradation	<i>Glutathione S-transferase family protein</i>
		Drug metabolism - cytochrome P450	<i>Glutathione S-transferase family protein</i>
		Drug metabolism - other enzymes	<i>Glutathione S-transferase family protein</i>
		Metabolism of xenobiotics by cytochrome P450	<i>Glutathione S-transferase family protein</i>
		Styrene degradation	<i>Glutathione S-transferase family protein</i>
	Sulfur metabolism	Sulfite oxidation	<i>MULTISPECIES: sulfite oxidase-like oxidoreductase</i>
Cellular processes	Membrane transport	Nuclear protein import	<i>Nuclear transport factor 2 family protein</i>
	Transporter	Cation/acetate symporter	<i>MULTISPECIES: cation acetate symporter</i>
	Adhesion	Binding to host proteins	WxL domain-containing protein

Abbreviations: **LB**, live birth. Regular font: LEfSe. Italic: DeSeq2. Bold: results from both analyses.

The NP group displayed an enrichment in genes associated with functions such as protein export, DNA replication and repair, and quorum sensing (Table 10).

Table 10. Bacterial functions and associated genes significantly enriched in NP.

Class	Function	Pathway	Gene
Cellular Processes	Cellular community-prokaryotes	Quorum sensing	Preprotein translocase subunit SecY
Environmental Information Processing	Membrane transport	Bacterial secretion system	Preprotein translocase subunit SecY
Metabolism	Nucleotide metabolism	Purine metabolism	DNA polymerase III subunit beta

Class	Function	Pathway	Gene
Genetic information processing	Folding, sorting, and degradation	Protein export	Preprotein translocase subunit SecY
	Replication and repair	DNA replication	DNA polymerase III subunit beta
		Homologous recombination	DNA polymerase III subunit beta
		Mismatch repair	DNA polymerase III subunit beta
Unknown	Unknown	Unknown	DUF340 domain-containing protein

Abbreviations: **NP**, no pregnancy. Genes come from LEfSe analysis.

Samples from the BP group had a high abundance of a wider range of gene functions, such as amino acid, carbohydrate, and xenobiotic metabolism. Additionally, a higher presence of genes related to metabolism of terpenoids and polyketides and replication was also observed (Table 11).

Table 11. Bacterial functions and associated genes significantly enriched in BP.

Class	Function	Pathway	Gene
Genetic Information Processing	Replication and repair	DNA helicase	ATP dependent helicase HrpB
Metabolism	Amino acid metabolism	Arginine and proline metabolism	Glutamate-5-semialdehyde dehydrogenase
			Gamma-glutamyl phosphate reductase
		Lysine biosynthesis	Dihydrodipicolinate synthase family protein
		Lysine degradation	MULTISPECIES: Acetyl-CoA acetyltransferase
		Tryptophan metabolism	MULTISPECIES: Acetyl-CoA acetyltransferase
		Valine, leucine, and isoleucine degradation	MULTISPECIES: Acetyl-CoA acetyltransferase
	Biosynthesis of other secondary metabolites	Phenazine biosynthesis	Propionyl-CoA synthetase
	Energy metabolism	Carbon fixation pathways in prokaryotes	MULTISPECIES: Acetyl-CoA acetyltransferase
			Propionyl-CoA synthetase
		Methane metabolism	Propionyl-CoA synthetase

Class	Function	Pathway	Gene
Metabolism	Carbohydrate metabolism	Butanoate metabolism	MULTISPECIES: Acetyl-CoA acetyltransferase
		Glyoxylate and dicarboxylate metabolism	MULTISPECIES: Acetyl-CoA acetyltransferase
		Glycolysis /Gluconeogenesis	Propionyl-CoA synthetase
		Propanoate metabolism	Propionyl-CoA synthetase
		Pyruvate metabolism	Propionyl-CoA synthetase
	Global metabolism	Biosynthesis of secondary metabolites	MULTISPECIES: Acetyl-CoA acetyltransferase
		Carbon metabolism	MULTISPECIES: Acetyl-CoA acetyltransferase
		Fatty acid metabolism	MULTISPECIES: Acetyl-CoA acetyltransferase
			NADPH-dependent 2,4-dienoyl-CoA reductase
		Microbial metabolism in diverse environments	MULTISPECIES: Acetyl-CoA acetyltransferase
	Lipid metabolism	Fatty acid degradation	MULTISPECIES: Acetyl-CoA acetyltransferase
		Fatty acid elongation	MULTISPECIES: Acetyl-CoA acetyltransferase
	Metabolism of terpenoids and polyketides	Terpenoid backbone biosynthesis	MULTISPECIES: Acetyl-CoA acetyltransferase
	Xenobiotics biodegradation and metabolism	Aminobenzoate degradation	Propionyl-CoA synthetase
			Propionyl-CoA synthetase
		Benzoate degradation	NADP dependent oxidoreductase
			MULTISPECIES: Acetyl-CoA acetyltransferase
	Polyketide metabolism	Synthesis of polyketides	Polyketide cyclase
Unknown	Unknown	Unknown	DUF2815 domain-containing protein
			DUF1768 domain-containing protein

Abbreviations: **BP**, biochemical pregnancy. Genes come from LEfSe analysis.

Among CM samples, metabolism of molecules such as amino acid, carbohydrate, and xenobiotic were also overrepresented. Moreover, protein folding and degradation, translation, and necrosis were found in higher abundance among CM samples, as

well as some others related to the endocrine system and metabolic diseases such as insulin/glucagon signaling pathway through glycogen/starch metabolism (Table 12).

Table 12. Bacterial functions and associated genes significantly enriched in CM.

Class	Function	Pathway	Gene
Cellular processes	Cell growth and death	Necroptosis	Glycogen/starch/Alpha-glucan family phosphorylase
	Cellular community-prokaryotes	Biofilm formation - Escherichia coli	Glycogen/starch/Alpha-glucan family phosphorylase
		Biofilm formation	Type VI secretion protein VskK
	RNA processing, cell cycle regulation	Ribosomal subunit biogenesis, protein binding	GTPase Era
Environmental Information Processing	Membrane transport	Bacterial secretion system	Type VI secretion protein VskK
Genetic information processing	Folding, sorting, and degradation	Proteasome	50S ribosomal protein L11
		Sulfur relay system	Alkyl hydroperoxide reductase
	Translation	Ribosome	50S ribosomal protein L11
			30S ribosomal protein S6
		Aminoacyl-tRNA biosynthesis	Isocitrate dehydrogenase (NADP)
			Threonine-tRNA ligase
		Ribosome biogenesis in eukaryotes	AAA family ATPase
	Replication and repair	DNA replication proteins	AAA family ATPase
	Transcription factors	Transcriptional repressor	Transcriptional regulator NrdR
Human diseases	Endocrine and metabolic disease	Insulin resistance	Glycogen/starch/Alpha-glucan family phosphorylase
Metabolism	Amino acid metabolism	Cysteine and methionine metabolism	Cystathionine gamma-synthase
		Phenylalanine metabolism	Fumarate reductase-quinol flavoprotein subunit
		Valine, leucine, and isoleucine degradation	Sugar ABC transporter substrate-binding protein
	Metabolism of other amino acids	Glutathione metabolism	Isocitrate dehydrogenase (NADP)
		Selenocompound metabolism	Cystathionine gamma-synthase

Class	Function	Pathway	Gene
Metabolism	Carbohydrate metabolism	Butanoate metabolism	Fumarate reductase-quinol flavoprotein subunit
		Butanoate metabolism	Sugar ABC transporter substrate-binding protein
		C5-Branched dibasic acid metabolism	Sugar ABC transporter substrate-binding protein
		Citrate cycle (TCA cycle)	Fumarate reductase-quinol flavoprotein subunit
			Isocitrate dehydrogenase (NADP)
			Acyl_CoA thioesterase II
		Fructose and mannose metabolism	Acyl_CoA thioesterase II
		Glycolysis /Gluconeogenesis	Acyl_CoA thioesterase II
		Pentose phosphate pathway	Acyl_CoA thioesterase II
		Pyruvate metabolism	Fumarate reductase-quinol flavoprotein subunit
		Starch and sucrose metabolism	Glycogen/starch/Alpha-glucan family phosphorylase
		Synthesis of glycoconjugates	1,3-beta-galactosyl-N-acetylhexosamine phosphorylase
	Energy metabolism	Carbon fixation pathways in prokaryotes	Isocitrate dehydrogenase (NADP)
			Fumarate reductase-quinol flavoprotein subunit
		Oxidative phosphorylation	Fumarate reductase-quinol flavoprotein subunit
		Sulfur metabolism	Cystathionine gamma-synthase
		Protein metabolism	Peptidase E
			Peptidase S1
		Iron acquisition	MULTISPECIES: ferrous iron transporter B
	Global metabolism	Biosynthesis of amino acids	Cystathionine gamma-synthase
			Isocitrate dehydrogenase (NADP)
		Biosynthesis of cofactors	Bifunctional folypolyglutamate synthase-dihydrofolate synthase
		Carbon metabolism	Isocitrate dehydrogenase (NADP)
		Microbial metabolism in diverse environments	Isocitrate dehydrogenase (NADP)
		2-Oxocarboxylic acid metabolism	Isocitrate dehydrogenase (NADP)

Class	Function	Pathway	Gene
Metabolism	Metabolism of cofactors and vitamins	Folate biosynthesis	Bifunctional folylpolyglutamate synthase-dihydrofolate synthase
	Xenobiotics biodegradation and metabolism	Steroid degradation	Fumarate reductase-quinol flavoprotein subunit
Organismal systems	Endocrine system	Glucagon signaling pathway	Glycogen/starch/Alpha-glucan family phosphorylase
		Insulin signaling pathway	Glycogen/starch/Alpha-glucan family phosphorylase
Unknown	Unknown	Unkown	DUF1541 domain-containing protein

Abbreviations: **CM**, clinical miscarriage. Genes in regular font come from LEfSe.

These findings reveal that the microbiome in different reproductive outcomes correlates with distinct bacterial functions, such as quorum sensing, secretion system and overall metabolism, suggesting bacterial communities in the endometria of patients with reproductive failure behave differently than those associated with reproductive success.

5.2.2 Functional profiling of bacterial taxa

When considering associations between bacterial genes and taxonomical data for each reproductive outcome, we encountered a few taxa-function associations among LB and NP (Figures 38 and 39), whilst more complex links between bacterial taxa and functions were observed in BP and CM groups (Figures 40 and 41). In LB we highlight the negative association between ABC (ATP-binding cassette) transporters with several *Lactobacillus* species, while some transposases and acyl-CoA dehydrogenase were positively associated with *L. crispatus* and *L. gallinarum* (Figure 38).

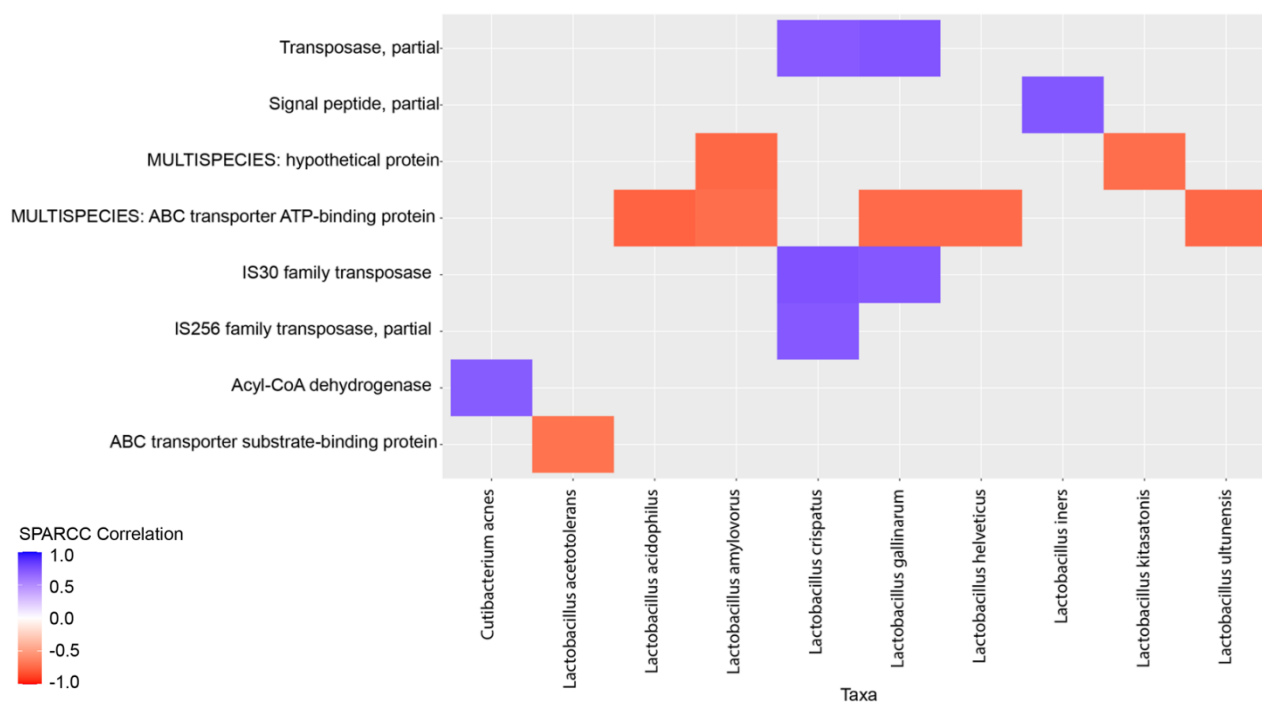


Figure 38. Coupled taxonomic and functional networks for the bacterial communities in the endometrium of infertile women from the LB group. SparCC (Sparse Correlations for Compositional data) networks considering the taxa and functions showed significant correlations found for each reproductive outcome (values >0.7 and <-0.7 ; p -value <0.05).

In samples from the NP group, main features include positive associations between transposases and lysin with *Lactobacillus* spp., including *L. crispatus* but not *L. iners*. Moreover, *Bifidobacterium* spp. and *Gardnerella* had unique associations, including positive links with alpha-amylase and 4-alpha-glucanotransferase, both involved in starch metabolism and biofilm formation. Lastly, *L. iners* positively associated with YSIRK/LPXTG domain surface protein, which may increase virulence by promoting adherence to host cells (Figure 39).

RESULTS

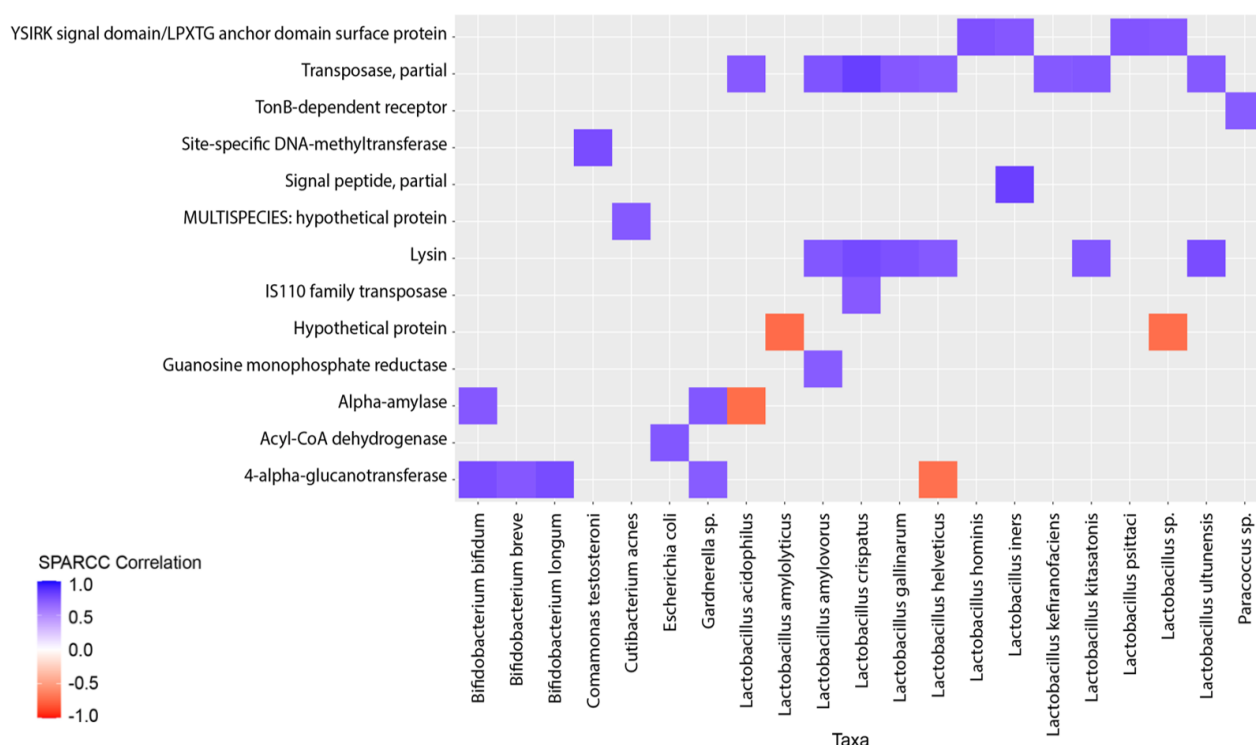


Figure 39. Coupled taxonomic and functional networks for the bacterial communities in the endometrium of infertile women from the NP group. SparCC (Sparse Correlations for Compositional data) networks considering the taxa and functions showed significant correlations found for each reproductive outcome (values >0.7 and <-0.7 ; p -value <0.05).

In the BP group, transposases once again associated with several taxa, either positively (*L. crispatus*, *L. helveticus*, and *L. amylophilus*, among other *Lactobacilli*) or negatively (*A. vaginae*, *B. bifidum*, *E. coli*, *G. vaginalis*, and *L. iners*). Lastly, MFS (Major Facilitator Superfamily) and ABC transporters mostly had negative associations with several commensals, except for a positive correlation with *L. crispatus* (Figure 40).

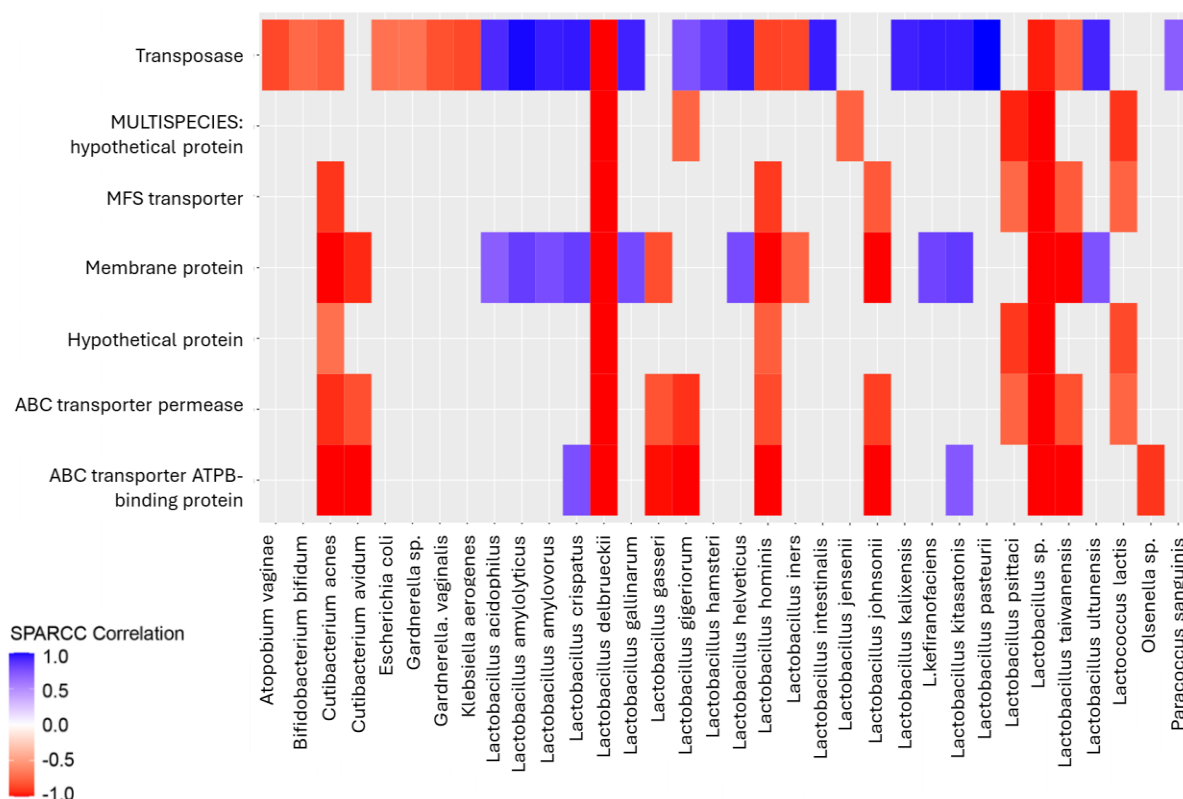


Figure 40. Coupled taxonomic and functional networks for the bacterial communities in the endometrium of infertile women from the BP group. SparCC (Sparse Correlations for Compositional data) networks considering the taxa and functions showed significant correlations found for each reproductive outcome (values >0.7 and <-0.7 ; p -value <0.05).

Lastly, the CM group was characterized by positive associations between *Bifidobacterium* spp., *G. vaginalis*, and *Olsenella* sp. with several genes from the ABC and MFS transporter superfamilies, sensor histidine kinase, and membrane protein. The opposite was observed for *Lactobacillus* spp., as several Lactobacilli negatively associated with these genes (Figure 41).

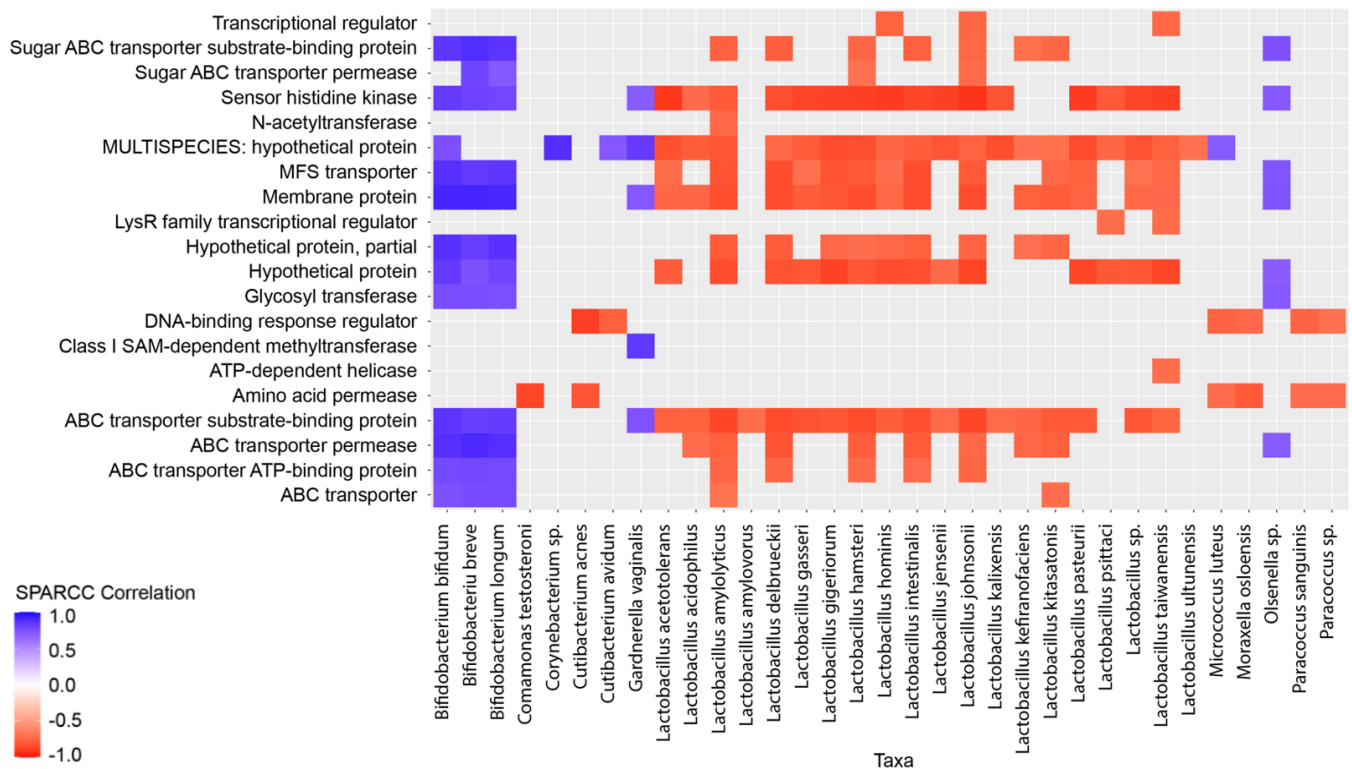


Figure 41. Coupled taxonomic and functional networks for the bacterial communities in the endometrium of infertile women from the CM group. SparCC (Sparse Correlations for Compositional data) networks considering the taxa and functions showed significant correlations found for each reproductive outcome (values >0.7 and <-0.7 ; p -value <0.05).

These data suggest that differences in the microbiota composition significantly associate with distinct functional bacterial networks, which could influence reproductive outcomes by changing the endometrial microenvironment and how it interacts with the developing embryo.

5.2.3 Virulence factors

To assess which genes found in the endometrial microbiome could assist the bacteria colonize the host at the cellular level, we used the Virulence Factor Database (VFDB) and observed a significant enrichment of 360 virulence factor-related genes in reproductive failure patients compared to those with an LB, whereas additional 800 genes were only found in reproductive failure-associated groups. Among the different

categories, adherence, effector delivery system, and immune modulation contribute with more than 500 genes that were absent in the LB group.

Compared to LB, the NP group displayed the enrichment of a total of 11 genes, whereas 412 other genes distributed into 14 categories that were not found in the reproductive success group (Table 13).

Table 13. Number of virulence factor-associated genes in the NP group compared to LB.

Category	Differentially abundant genes	Uniquely found genes
Adherence	0	84
Antimicrobial activity/Competitive advantage	0	1
Biofilm	0	16
Effector delivery system	1	121
Exoenzyme	0	5
Exotoxin	3	14
Immune modulation	5	86
Invasion	0	3
Motility	0	36
Nutritional/Metabolic factor	0	33
Others	0	9
Post-translational modification	0	3
Regulation	1	0
Species-associated genes	0	0
Stress survival	1	1

Left: differentially abundant genes in comparison to the LB group. Right: genes completely absent in the LB group.

Regarding differentially abundant genes, most of them were related to: effector delivery system ($p\text{-adj}_{lirB}=0.03$); genes encoding exotoxins (pneumolysin, $p\text{-adj}_{ply}=0.035$; and hemolysin, $p\text{-adj}_{sly}=0.02$; and $p\text{-adj}_{HMPREF0424_RS00505}=0.03$); immune modulation ($p\text{-adj}_{wbjD/wecB}=0.03$; $p\text{-adj}_{wecB}=0.04$; and $p\text{-adj}_{wlaN}=0.03$, etc); regulation (histidine kinase involved in response to environment, $p\text{-adj}_{senX3}=0.03$), and stress

survival against acute toxicity of bile salts ($p\text{-adj}_{bsh}=0.03$) (Annex 1, Figure 42 and Table 14).

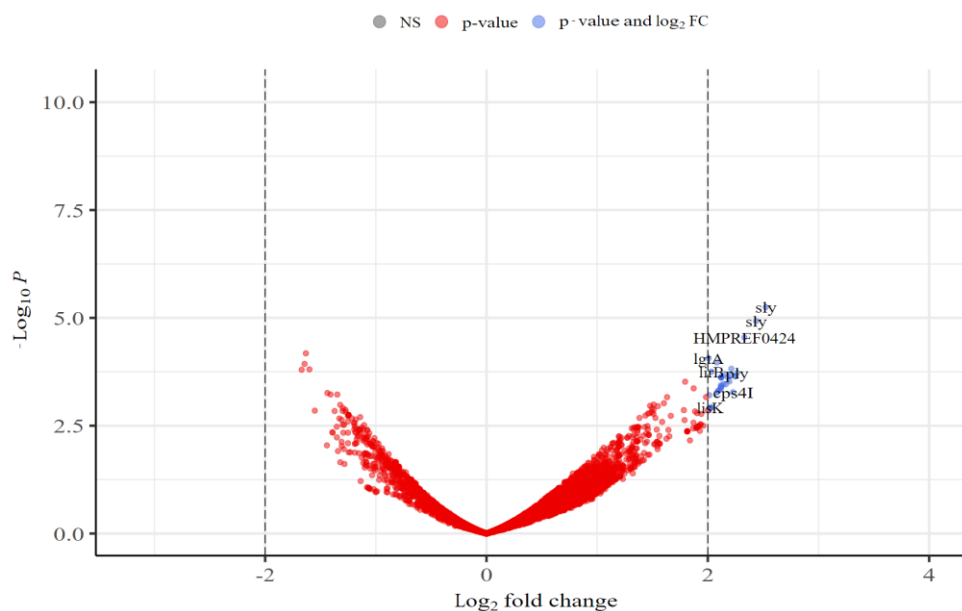


Figure 42. Virulence factors genes enriched in NP compared to LB. Volcano plot depicts virulence factor genes with higher abundance among patients with NP ($\log_2$ fold change ≥ 2.0 and ≤ -2 , $p\text{-adj.} \leq 0.05$).

Table 14. Virulence factors differentially abundant in NP compared to LB.

Category	VF gene name	Description	BaseMean	log2FC	p-adj
Effector delivery system (VFC0086)	lirB	Dot/lcm type IV secretion system effector LirB Dot/lcm T4SS secreted effectors (VF0798)	11.27	2.03	0.03
Exotoxin (VFC0235)	ply	Pneumolysin (VF0149)	11.78	2.26	0.03
	HMPREF0424_RS00505	thiol-activated cytolysin family protein Vaginolysin (VLY) (VF1316)	6.58	2.33	0.03
	sly	hemolysin Suilysin (VF1062)	5.75	2.43	0.02
Immune modulation (VFC0258)	lgtA	N-acetylglucosamine glycosyltransferase LOS (VF0044)	2.91	2.00	0.03
	wlaN	beta-1;3 galactosyltransferase LOS (VF0326)	4.13	2.08	0.03
	BCE_RS25805	LytR family transcriptional regulator Polysaccharide capsule (VF0659)	13.75	2.09	0.05

Category	VF gene name	Description	BaseMean	log2FC	p-adj
Immune modulation (VFC0258)	wecB	UDP-N-acetylglucosamine 2-epimerase (non-hydrolyzing) LPS (VF0085)	12.96	2.19	0.04
	wbjD/wecB	UDP-N-acetylglucosamine 2-epimerase (non-hydrolyzing) Capsular polysaccharide (VF0624)	13.44	2.24	0.03
Regulation (VFC0301)	senX3	two-component sensor histidine kinase SenX3 (VF0857)	13.60	2.09	0.04
Stress survival (VFC0282)	bsh	bile salt hydrolase BSH (VF0350)	9.46	2.16	0.04

Abbreviations: **FC**, fold change. For full version of table, please check Annex 1.

Regarding genes that were not present in the LB group, most of them belong to adherence, effector delivery system, and immune modulation categories (Table 13 and Annex 2). For instance, many genes were associated with formation of pili and adhesive fimbriae, T3 and T4SS structures, and biosynthesis of capsule and surface proteins (Annex 2).

We found no significant differences in the abundance of virulence factors when comparing BP and LB groups. Nonetheless, 125 genes that were present in the BP group were completely absent in samples coming from patients who had a LB (Table 15). Among these, again there was an overrepresentation of genes associated with adherence (such as those involved in pili formation, like pil family, *tapM*, and *sapA*), effector delivery system from types III and IV (such as *bapB*, *cdsC*, *dotU*, and *hcp1*), and immune modulation (*htrB*, *mce4D*, and *gtrA*); among other less representative categories (Annex 3).

Table 15. Number of virulence factor-associated genes in the BP group.

Category	Differentially abundant genes	Uniquely found genes
Adherence	0	23
Biofilm	0	4
Effector delivery system	0	38
Exoenzyme	0	6
Exotoxin	0	5
Immune modulation	0	28
Motility	0	8
Nutritional/Metabolic factor	0	10
Others	0	1
Regulation	0	1
Stress survival	0	1

Left: differentially abundant genes in comparison to the LB group. Right: genes completely absent in the LB group.

Meanwhile, the CM group displayed enrichment of 349 genes compared to the LB group, and additional 263 genes that were only found in this group (Table 16).

Table 16. Number of virulence factor-associated genes in the CM group.

Category	Differentially abundant genes	Uniquely found genes
Adherence	45	62
Antimicrobial activity/Competitive advantage	1	0
Biofilm	7	6
Effector delivery system	28	77
Exoenzyme	8	5
Exotoxin	13	5
Immune modulation	165	41
Invasion	3	2
Motility	8	19
Nutritional/Metabolic factor	50	30
Others	3	10
Post-translational modification	3	0
Regulation	8	3
Species-associated genes	1	0
Stress survival	6	3

Left: differentially abundant genes in comparison to the LB group. Right: genes completely absent in the LB group.

Among those, we highlight genes related to effector delivery system, most of them related to ATPase secretion systems of different types; production of exotoxins like hemolysins ($p\text{-adj}_{hlyIII}=0.00$; and $p\text{-adj}_{hlyA}=0.00$); and hemolysins/cytolysins ($p\text{-adj}_{cylG}=0.03$; and $p\text{-adj}_{cylZ}=0.03$), immune modulation genes, regulation including the *senX3* histidine kinase, also found enriched in NP ($p\text{-adj}_{senX3}=0.00$) and the two-component signal transduction system PhoP-PhoR that enables bacteria to adapt to environmental stimuli ($p\text{-adj}_{phoP}=0.00$; and $p\text{-adj}_{phoR}=0.00$), stress survival with genes related to proteasome ATPases ($p\text{-adj}_{mpa}=0.00$), the bile salt hydrolase also significant in NP group ($p\text{-adj}_{bsh}=0.00$), and the ABC transporter ($p\text{-adj}_{mntA}=0.01$) (Figure 43 and Annex 4).

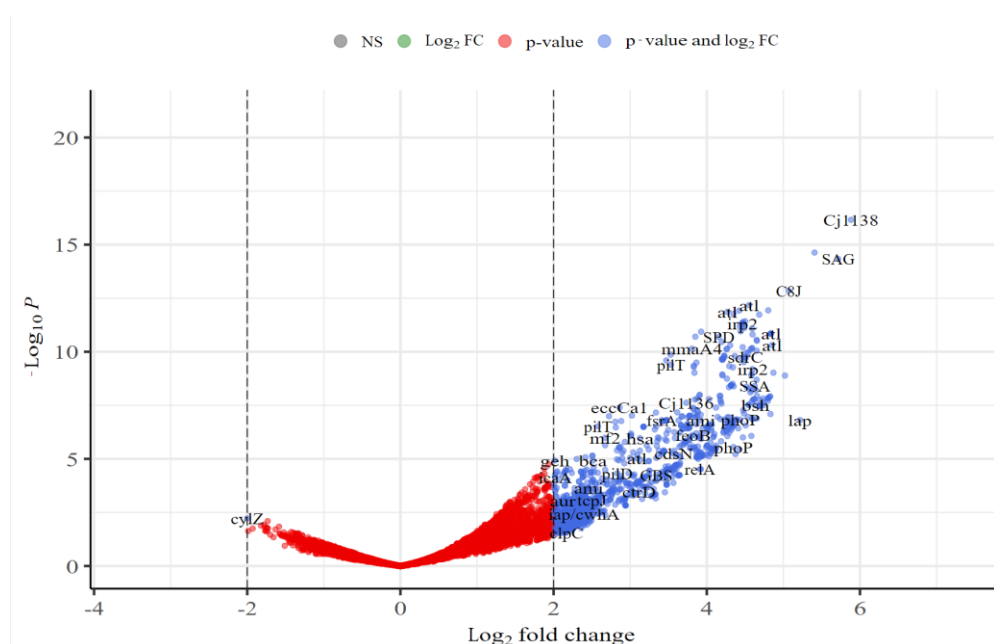


Figure 43. Virulence factors genes enriched in CM compared to LB. Volcano plot depicts virulence factor genes with higher abundance among patients with CM ($\log_2$ fold change ≥ 2.0 and ≤ -2 , $p\text{-adj.} \leq 0.05$).

Besides, the CM group also displayed the differential expression of genes associated with adhesion factors, antimicrobial activity, biofilm formation, exoenzymes, invasion, motility, nutritional factors, post-translational modifications,

and one Enterococcal surface protein *esp* from *E. faecalis* (Figure 43 and Annex 4). Among them, it is important to highlight the significant enrichment of adhesion-related genes involved in pili function (members A, B, D, H and T of the *pil* family), and the type-II secretion system protein *GspE* found in Enterobacteriaceae; adhesins including the *sdr* family composed of Ser-Asp rich fibrinogen-binding protein, and the members of the fibronectin-binding protein (*fpb*) family. Biofilm-related genes with the highest increase in CM included genes contributing to the formation and development of alginate and other polysaccharide matrices that help bacterial survival against environmental stress and host responses (*algW*, *mucD*, and *pgaC*). Interestingly, several genes were significantly enriched in CM involving the transport of ions like Mg^{2+} (*mgtB* and *mgtC*), Mn^{2+} (*ctpC*) and iron metabolism including siderophores, iron-regulated ABC transporters, and heme uptake systems (Figure 43 and Annex 4). Regarding the 263 genes that were not found in the LB group, most of them associated with adherence, effector delivery system, and immune modulation categories (Table 16). Additionally, several genes were related to other functions such as motility (flagella structure and function), biofilm formation, and nutritional factors (iron transporters and Heme uptake systems) (Annex 5).

These results suggest that a dysbiotic bacterial community leads to the higher presence of various virulence factor-related genes (mainly those related to adherence, effector delivery system, and immune modulation) in patients with reproductive failure following assisted reproductive technology treatment, which could explain part of the pathogenesis of infertility.

5.2.4 Antibiotic resistance

Building up on our previous findings that indicated an association between reproductive outcomes and genes from the ABC and MFS superfamilies involved in the transport of molecules such as sugars, ions, and antibiotics – with potential implications in antibiotic resistance –, we next investigated the presence of antibiotic resistance-related genes within our samples.

Using the CARD database, we observed an enrichment in the presence of antibiotic resistance-related genes in reproductive failure patients. When comparing the abundance of antibiotic resistance-associated genes in samples from the NP group, we discovered the significantly increased abundance of the *IsaC* gene which encodes an ABC transporter and confers resistance to lincosamides in the NP group ($p\text{-adj}_{IsaC}=0.04$) (Table 17).

Table 17. Antibiotic resistance-related genes differentially abundant in NP compared to LB.

Resistance Mechanism	Drug Class	AMR gene family	ARO Term	log2FC	p-adj
antibiotic efflux	lincosamide	Isa-type ABC-F protein	IsaC	2.69	0.04

Abbreviations: **AMR**, antimicrobial resistance; **ARO**, antibiotic resistance ontology; **LB**, live birth; **NP**, no pregnancy.

Despite significant differences in gene abundance, we also observed 33 antibiotic resistance-associated bacterial genes in NP that were absent in samples from the LB group. Although most of them act through antibiotic efflux via ABC, RND (resistance-nodulation-cell division), or MFS transporters, inactivation of antibiotics through lactamases, target alteration, protection, or replacement were also observed (Table 18).

Table 18. Antibiotic resistance-related genes present only in samples from NP group compared to LB.

Resistance Mechanism	Drug Class	AMR gene family	ARO Term
antibiotic efflux	aminoglycoside	resistance-nodulation-cell division (RND) antibiotic efflux pump	<i>smeA</i>
			<i>smeS</i>
	disinfecting agents and antiseptics	major facilitator superfamily (MFS) antibiotic efflux pump	<i>qacE</i>
			<i>qacA</i>
	fluoroquinolone	resistance-nodulation-cell division (RND) antibiotic efflux pump	<i>AcrS</i>
			<i>adeF</i>
	lincosamide	ATP-binding cassette (ABC) antibiotic efflux pump	<i>lmrD</i>
	macrolide	resistance-nodulation-cell division (RND) antibiotic efflux pump	<i>MexW</i>
			<i>Pseudomonas aeruginosa CpxR</i>
			<i>smeD</i>
antibiotic inactivation	aminoglycoside	aminoglycoside bifunctional resistance protein	<i>AAC(6')-Ie-APH(2'')-Ia bifunctional protein</i>
			<i>APH(3')-Ia</i>
			<i>APH(6)-Id</i>
	carbapenem	IMI beta-lactamase	<i>IMI-5</i>
		OXA beta-lactamase	<i>OXA-53</i>
		Rm3 family beta-lactamase	<i>Rm3</i>
	cephalosporin	L1 family beta-lactamase	<i>L1 beta-lactamase</i>
	cephamycin	CfxA beta-lactamase	<i>CfxA2</i>
			<i>CfxA3</i>
	macrolide	mgt macrolide glycotransferase	<i>mgtA</i>
	monobactam	TEM beta-lactamase	<i>TEM-126</i>
			<i>TEM-192</i>
	rifamycin	rifampin monooxygenase	<i>iri</i>
		rifampin phosphotransferase	<i>rphA</i>
antibiotic target alteration	macrolide	Erm 23S ribosomal RNA methyltransferase	<i>ErmB</i>
	peptide	pmr phosphoethanolamine transferase	<i>ArnT</i>
antibiotic target protection	tetracycline	tetracycline-resistant ribosomal protection protein	<i>tet(32)</i>
			<i>tet(O/32/O)</i>
			<i>tet(O/M/O)</i>
antibiotic target replacement	diaminopyrimidine	trimethoprim resistant dihydrofolate reductase dfr	<i>dfrC</i>

Abbreviations: **AMR**, antimicrobial resistance; **ARO**, antibiotic resistance ontology; **NP**: no pregnancy; **LB**: live birth.

No differentially abundant genes were found when comparing BP and LB groups. Nonetheless, we observed eight genes that were completely absent in samples coming from patients who had a live birth. Again, genes coding for ABC, RND and MFS transporters, alongside beta-lactamases, were the most representative in this group (Table 19).

Table 19. Antibiotic resistance-related genes present only in samples from the BP group compared to LB.

Resistance Mechanism	Drug Class	AMR gene family	ARO Term
antibiotic efflux	macrolide	RND antibiotic efflux pump	<i>MexB</i>
	nitroimidazole	ABC antibiotic efflux pump	<i>msbA</i>
	tetracycline	MFS antibiotic efflux pump	<i>emrK</i>
antibiotic inactivation	carbapenem	CRD3 beta-lactamase	<i>CRD3-1</i>
		OXA beta-lactamase	<i>OXA-457</i>
		OXA beta-lactamase	<i>OXA-681</i>
antibiotic target alteration	macrolide	Erm 23S ribosomal RNA methyltransferase	<i>Erm(33)</i>
	peptide	pmr phosphoethanolamine transferase	<i>eptA</i>

Abbreviations: **AMR**, antimicrobial resistance; **ARO**, antibiotic resistance ontology; **BP**: biochemical pregnancy; **LB**: live birth.

Lastly, when comparing CM and LB groups, increased abundance of the *rpoB* gene (rifamycin-resistant beta-subunit of RNA polymerase) was found among patients with a clinical miscarriage ($p\text{-adj}_{rpoB}=0.01$) (Table 20).

Table 20. Antibiotic resistance-related genes differentially abundant in CM compared to LB.

Resistance Mechanism	Drug Class	AMR gene family	ARO Term	log2FC	p-adj
antibiotic target alteration	rifamycin	rifamycin-resistant beta-subunit of RNA polymerase (<i>rpoB</i>)	<i>Bifidobacterium adolescentis rpoB</i> mutants conferring resistance to rifampicin	2.77	0.01

Abbreviations: **AMR**, antimicrobial resistance; **ARO**, antibiotic resistance ontology; **CM**, clinical miscarriage; **LB**, live birth.

Despite significant differences in gene abundance, we also observed 16 antibiotic resistance-associated bacterial genes that were only present in the CM group. Most of them act through antibiotic efflux via ABC, RND, or MFS transporters, but genes associated with inactivation of antibiotics, target alteration, and target protection were also observed (Table 21).

Table 21. Antibiotic resistance-related genes present only in samples from the CM group compared to LB.

Resistance Mechanism	Drug Class	AMR gene family	ARO Term
antibiotic efflux	fluoroquinolone	RND antibiotic efflux pump	<i>ceoB</i>
			<i>Escherichia coli acrA</i>
			<i>Klebsiella pneumoniae acrA</i>
	macrolide	RND antibiotic efflux pump	<i>MexK</i>
	nucleoside	MFS antibiotic efflux pump	<i>mdtN</i>
	phenicol	MFS antibiotic efflux pump	<i>cmx</i>
	other	MFS antibiotic efflux pump	H-NS
		ABC antibiotic efflux pump	<i>mgrA</i>
antibiotic inactivation	aminoglycoside	APH(3")	APH(3")-Ib
	cephalosporin	ADC beta-lactamase without carbapenemase activity	ADC-15
	lincosamide	lincosamide nucleotidyltransferase (LNU)	<i>InuA</i>
	macrolide	<i>gimA</i> family macrolide glycosyltransferase	<i>gimA</i>
	monobactam	TEM beta-lactamase	TEM-187
antibiotic target alteration	mupirocin-like	antibiotic-resistant isoleucyl-tRNA synthetase (ileS)	<i>Staphylococcus aureus</i> <i>mupA</i> conferring resistance to mupirocin
	peptide	<i>pmr</i> phosphoethanolamine transferase	<i>arnA</i>
antibiotic target protection	macrolide	Miscellaneous ABC-F subfamily ATP-binding cassette ribosomal protection proteins	<i>oleB</i>

Abbreviations: **AMR**, antimicrobial resistance; **ARO**, antibiotic resistance ontology; **CM**: clinical miscarriage; **LB**: live birth.

Interestingly, we observed the unique presence of some bacterial antibiotic resistance-associated genes commonly expressed by chronic endometritis-causing pathogens in reproductive failure patients. Examples include genes related to antibiotic target protection, efflux, and target alteration from *Escherichia coli* (e.g. *tet(32)*, *emrK*, and *eptA*); *Klebsiella pneumoniae* (e.g. *acrA*, *qacE*, and *ArnT*); and *Staphylococcus aureus* (e.g. *qacA*, *tet(K)*, and *mupA*) (Tables 18, 19, and 21).

These results describe the significant enrichment or unique detection of genes related to resistance against various antibiotic classes in NP, BP and CM compared to the LB group, mainly those genes associated with beta-lactamases or MFS, RND, and ABC antibiotic efflux pumps.

5.3 TRANSCRIPTOMICS IN THE DYSBIOTIC ENDOMETRIUM

Considering changes in the microbiome may lead to differential host gene expression, we next conducted a metatranscriptomic (MTX) analysis of endometrial tissue to evaluate the host and microbial gene expression within the uterine cavity of our set of patients.

After RNA extraction, we obtained an average of 265.34 ng/μL of RNA per sample (range 14.7 ng/μL – 690.8 ng/μL). Library preparation yielded an average of 5.15 ng/μL (range 1.02 ng/μL – 15.3 ng/μL) with an average size of 377 bp of genetic material, specifically the transcriptome of the samples. Sequencing analysis of EB samples using MTX yielded 1,878,809,627 paired reads. Applying quality control parameters provided a mean of 32,961,572 joined sequencing reads per sample (range 20,175,101 – 65,341,458). Importantly, from all reads obtained, 916,318,756 (around 49%) were human, whilst the other 962,490,870 (51%) belonged to the microbiota. Further bioinformatics analysis cleared an average of 56% of bacterial reads per

sample (range 52 – 65%) due to their ribosomal origin or unannotated bacterial genes. As a result of the very low amount of remaining bacterial reads and the consequent low reading depth, microbial transcriptomic analysis could not be conducted on this study. Therefore, all subsequent results presented in this section are derived from the host transcriptome.

Transcriptomic analysis using bulk RNA-seq of endometrial biopsies revealed 90 differentially expressed human genes ($\log_2$ fold change ≥ 2.0 and ≤ -2.0 , $FDR \leq 0.05$) in patients who suffered from reproductive failure compared to patients who had a live birth (Tables 22-24).

The NP group presented 74 differentially expressed genes, with 73 downregulated and 1 upregulated genes compared to the LB group. The immune system, responsible for orchestrating the symbiotic balance between the microbiota and the host, became significantly deranged in NP endometria through the downregulation of several genes related to innate and adaptative immunity and cytokine signaling (Figure 44 and Table 22).

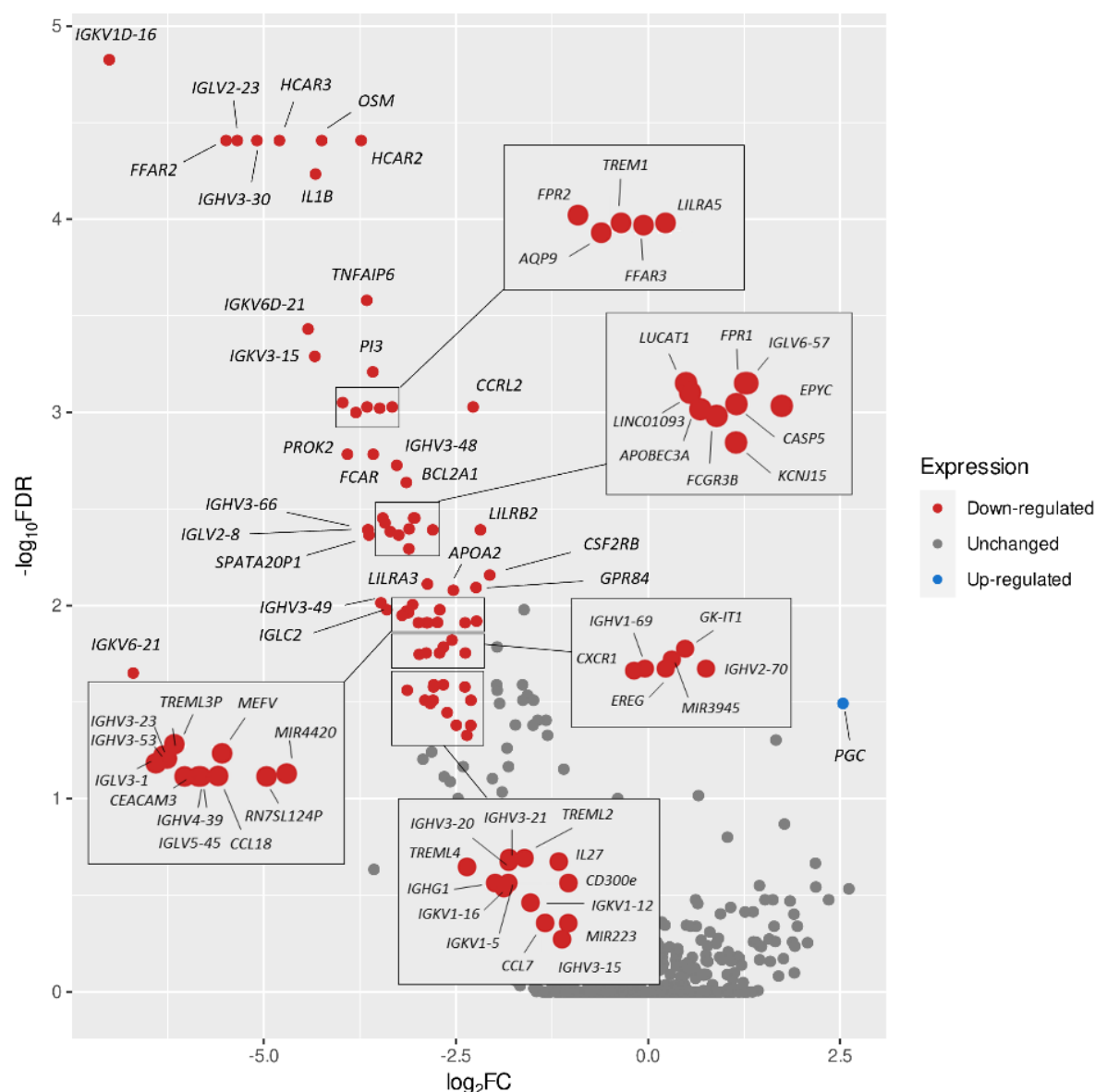


Figure 44. Differential gene expression of endometrial samples from the NP group. Volcano plot demonstrates several genes differentially expressed ($\log_2$ fold change ≥ 2.0 and ≤ -2.0 , $\text{FDR} \leq 0.05$) within the NP group when compared to LB. Abbreviations: **LB**, live birth; **NP**, no pregnancy.

Among these, genes with most substantial changes in expression were related to the humoral immune response, such as the IGH, IGL, and IGK families, which encode for heavy and light chains of immunoglobulins (for instance, *IGHV3-30* had a 34-fold decrease in expression compared to LB, whilst *IGKV1D-16* and *IGKV6-21* reached a 100-fold decrease in gene expression) (Table 22). Additionally, other downregulated

genes in the NP group were associated with cell-mediated immune response, such as chemokines and interleukins (*IL1B*, *IL27*, *CCL18*, *CCL7*, with fold decreases ranging from 5.6 to 13.4).

Table 22. Biological processes significantly up- and downregulated in no pregnancy compared to live birth.

Function	Contributing Gene	Log2 FC	FDR
<i>Down regulated in NP</i>			
Adaptative immune response	<i>IGKV1D-16</i>	-7.00	0.00
	<i>IGKV6-21</i>	-6.69	0.02
	<i>IGLV2-23</i>	-5.34	0.00
	<i>IGHV3-30</i>	-5.11	0.00
	<i>IGKV6D-21</i>	-4.42	0.00
	<i>IGKV3-15</i>	-4.33	0.00
	<i>IGHV3-66</i>	-3.64	0.00
	<i>IGLV2-8</i>	-3.63	0.00
	<i>IGHV3-49</i>	-3.47	0.00
	<i>IGLC2</i>	-3.40	0.00
	<i>IGHV3-48</i>	-3.26	0.00
	<i>IGLV3-1</i>	-3.19	0.01
	<i>IGHV3-23</i>	-3.14	0.01
	<i>IGHV3-53</i>	-3.11	0.01
	<i>IGLV6-57</i>	-3.03	0.00
	<i>IGHV1-69</i>	-2.88	0.01
	<i>IGLV5-45</i>	-2.88	0.01
	<i>IGHV4-39</i>	-2.86	0.01
	<i>IGKV1-16</i>	-2.83	0.03
	<i>IGKV1-5</i>	-2.79	0.02
	<i>IGHV3-21</i>	-2.79	0.02
	<i>IGHV3-20</i>	-2.79	0.02
	<i>IGKV1-12</i>	-2.61	0.03
	<i>IGHV2-70</i>	-2.37	0.01
	<i>IGHV3-15</i>	-2.35	0.04
	TREML2	-2.67	0.02
Innate immune response	<i>TREM1</i>	-3.67	0.00
	<i>LILRA5</i>	-3.34	0.00
	TREML4	-3.14	0.02
	FPR2	-3.99	0.00
	<i>IGHG1</i>	-2.91	0.02
	<i>FCAR</i>	-3.59	0.00
	<i>TREML3P</i>	-3.07	0.00

Function	Contributing Gene	Log2 FC	FDR
Cytokine-mediated immune response	<i>LILRA3</i>	-2.88	0.00
	<i>EREG</i>	-2.72	0.01
	<i>CSF2RB</i>	-2.08	0.00
Regulation of immune response	<i>FCGR3B</i>	-3.26	0.00
	<i>GPR84</i>	-2.25	0.00
	<i>HCAR3</i>	-4.81	0.00
	<i>CEACAM3</i>	-3.00	0.01
Inflammatory response	<i>FFAR2</i>	-5.48	0.00
	<i>OSM</i>	-4.35	0.00
	<i>IL1B</i>	-3.75	0.00
	<i>FFAR3</i>	-3.49	0.00
	<i>MEFV</i>	-2.72	0.00
	<i>CCL7</i>	-2.49	0.04
	<i>CXCR1</i>	-2.99	0.01
	<i>CCL18</i>	-2.74	0.01
	<i>CCRL2</i>	-2.29	0.00
	<i>TNFAIP6</i>	-3.68	0.00
Cellular defense response	<i>IL27</i>	-2.38	0.02
	<i>CD300E</i>	-2.32	0.02
	<i>LILRB2</i>	-2.20	0.00
	<i>FPR1</i>	-3.07	0.00
Antimicrobial activity	<i>APOBEC3A</i>	-3.37	0.00
	<i>AQP9</i>	-3.82	0.00
	<i>PI3</i>	-3.59	0.00
Post-transcriptional regulation of gene expression	<i>MIR3945</i>	-2.66	0.01
	<i>MIR223</i>	-2.32	0.04
	<i>MIR4420</i>	-2.23	0.01
Regulation of apoptosis	<i>BCL2A1</i>	-3.16	0.00
	<i>CASP5</i>	-3.122	0.00
	<i>HCAR2</i>	-4.26	0.00
Other	<i>PROK2</i>	-3.93	0.00
	<i>SPATA20P1</i>	-3.64	0.00
	<i>LUCAT1</i>	-3.47	0.00
	<i>LINC01093</i>	-3.42	0.00
	<i>KCNJ15</i>	-3.13	0.00
	<i>EPYC</i>	-2.79	0.00
	<i>GK-IT1</i>	-2.55	0.01
	<i>APOA2</i>	-2.53	0.00
	<i>RN7SL124P</i>	-2.38	0.01
Upregulated in NP			
Other	<i>PGC</i>	2.54	0.03

Abbreviations: **FC**, fold change; **FDR**, false discovery rate; **LB**, live birth; **NP**, no pregnancy. GPCRs are highlighted in bold.

Moreover, we observed a notable downregulation in G protein-coupled receptor signaling, such as free fatty acids receptors *FFAR2*, *FFAR3*, and *GPR84* (decreased by 44.9, 11.2 and 4.7-fold, respectively), hydroxycarboxylic acid receptors *HCAR2* and *HCAR3* (expressed 19 and 28 times less), formylpeptide receptor *FPR1* (exhibiting a 8.4-fold decrease), and chemokine receptors *CCRL2* and *CXCR1* (with fold decreases of 4.8 and 7.9, respectively) (Table 22). Noteworthy, other gene functions with relatively lower representation include post-transcriptional regulation of gene expression through miRNA such as *MIR3945*, *MIR223*, and *MIR4420* (with fold decreases ranging from 4.6 to 6.3), regulation and execution of apoptosis impacted through both *BCL2A1* and *CASP5* (characterized by a significant decrease in gene expression by at least 8.7-fold in each), ion channels such as *KCNJ15* (8.7-fold decrease), among others (Table 22).

Interestingly, downregulation of immune genes was not so representative among the other groups. The BP displayed the upregulation of only 13 genes compared to the LB group, but with a wider range of gene functions such as membrane trafficking (*CPNE7*, 4.5-fold increase), transcription regulation (*TFAP2A*, 4.3-fold increase), protein translation (*RPL17P22*, 24.3-fold increase), cell death (*CLU*, expressed 8.2 times more), complement system (*C8B*, increased by 15.1-fold), and GPCR signaling (angiotensin receptor *AGTR2*, with an increased gene expression of at least 64-fold) (Figure 45 and Table 23).

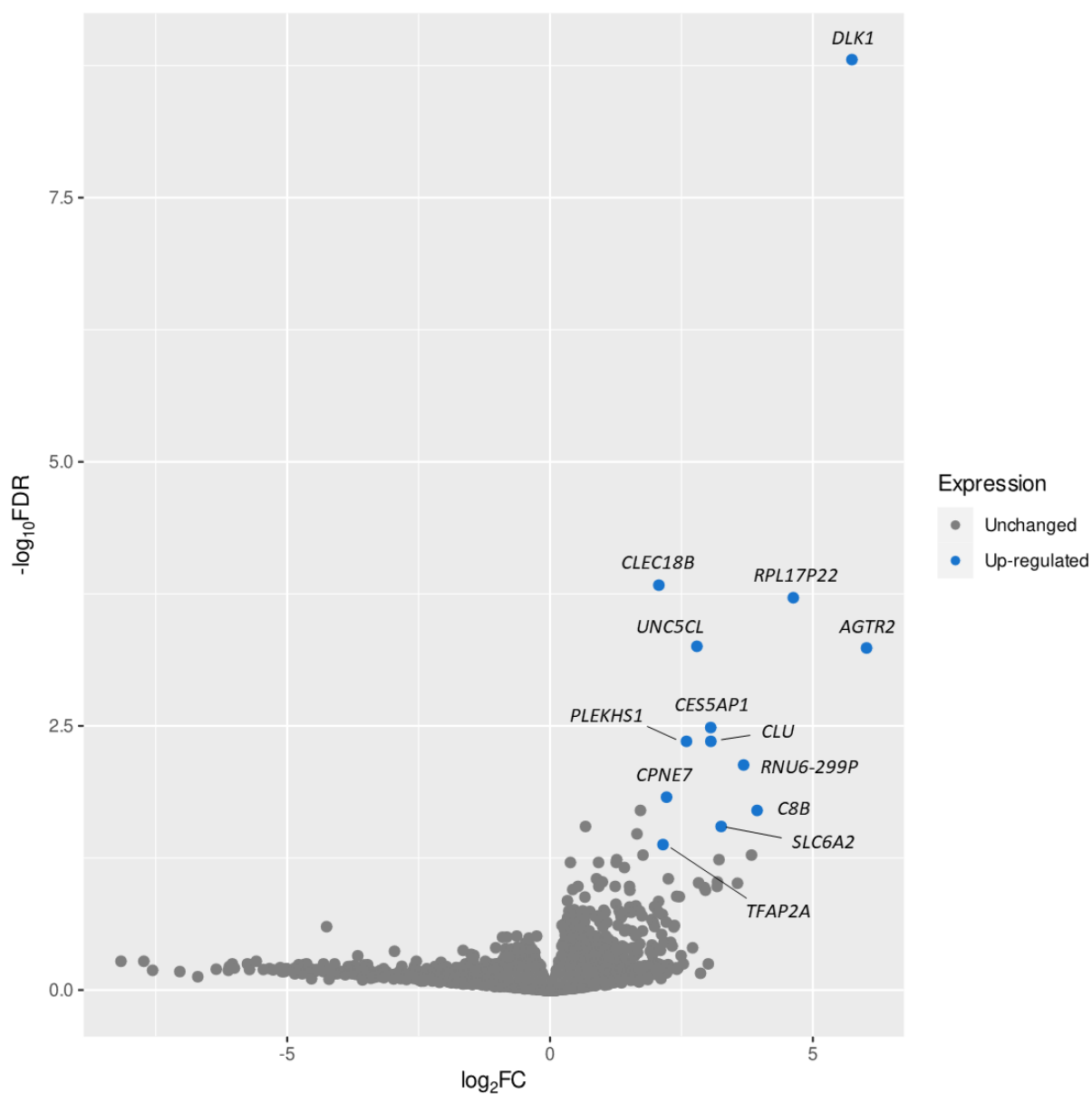


Figure 45. Differential gene expression of endometrial samples from the BP group. Volcano plot demonstrates several genes differentially expressed (log₂ fold change ≥ 2.0 and ≤ -2.0 , FDR ≤ 0.05) within the NP group when compared to LB. Abbreviations: **BP**, biochemical pregnancy; **LB**, live birth.

Table 23. Biological processes significantly upregulated in biochemical pregnancy compared to live birth.

Function	Contributing Gene	Log2 FC	FDR
Innate immune response	<i>C8B</i>	3.91	0.02
Inflammatory response	<i>AGTR2</i>	6.00	0.00
Host-pathogen and cell-cell interactions	<i>CLEC18B</i>	2.05	0.00
Regulation of cell growth	<i>DLK1</i>	5.73	0.00
Cell death and tumor progression	<i>CLU</i>	3.03	0.00
Transcriptional regulation	<i>TFAP2A</i>	2.12	0.04
Protein translation	<i>RPL17P22</i>	4.60	0.00
Transmembrane transport	<i>SLC6A2</i>	3.23	0.02
Signal transduction	<i>UNC5CL</i>	2.77	0.00
Transmembrane transport	<i>CPNE7</i>	2.19	0.01
Other	<i>RNU6-299P</i>	3.66	0.00
	<i>CES5AP1</i>	3.03	0.00
	<i>PLEKHS1</i>	2.57	0.00

Abbreviations: **BP**, biochemical pregnancy; **FC**, fold change; **FDR**, false discovery rate; **LB**, live birth. GPCR highlighted in bold.

Lastly, the CM group displayed the upregulation of 3 genes related to heavy chains of immunoglobulins (*IGHV3-73*, 7.5-fold increase), GPCR signaling (glucagon receptor *GLP2R*, 9.5-fold increase), and a third one with unknown function to this day (*FAM230A*, characterized by an increase in gene expression by at least 37-fold) (Figure 46 and Table 24).

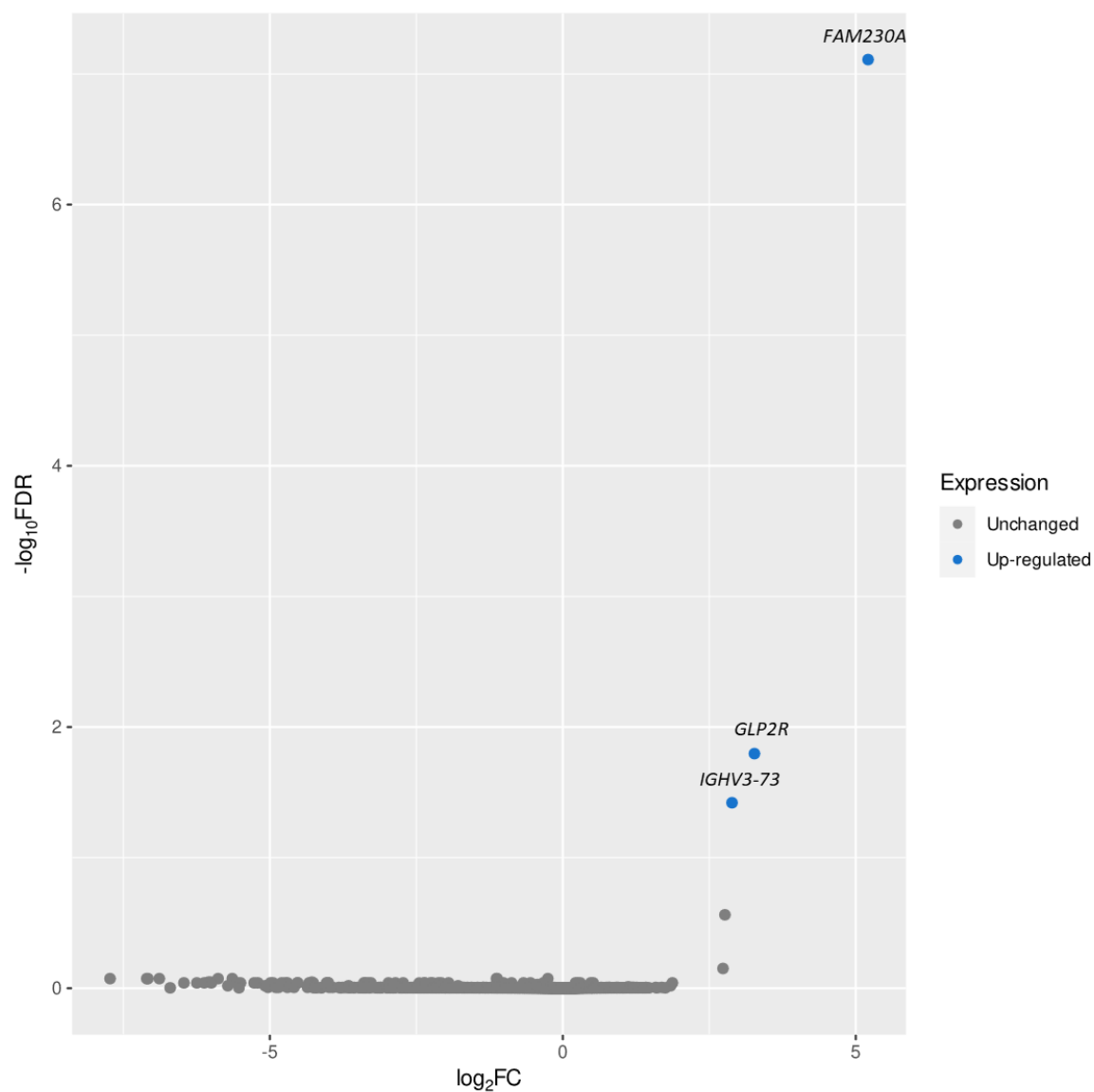


Figure 46. Differential gene expression of endometrial samples from the CM group. Volcano plot demonstrates several genes differentially expressed ($\log_2$ fold change ≥ 2.0 and ≤ -2.0 , $FDR \leq 0.05$) within the NP group when compared to LB. Abbreviations: **CM**, clinical miscarriage; **LB**, live birth.

Table 24. Biological processes significantly upregulated in clinical miscarriage compared to live birth.

Function	Contributing Genes	Log2 FC	FDR
Activation of immune response, defense response to other organism, and phagocytosis	<i>IGHV3-73</i>	2.91	0.03
Glucagon signaling	<i>GLP2R</i>	3.26	0.02
Other	<i>FAM230A</i>	5.22	0.00

Abbreviations: **CM**, clinical miscarriage; **FC**, fold change; **FDR**, false discovery rate; **LB**, live birth. GPCRs are highlighted in bold.

Ultimately, these results demonstrate different signaling pathways that are either up or downregulated in each of the subgroups of reproductive failure, which sheds light on the molecular mechanisms underlying infertility. Their association with fluctuations in the endometrial microbiota composition remain to be unraveled.

6. DISCUSSION

Although an altered endometrial microbiota composition has been associated with adverse pregnancy outcomes and fertility-related diseases (Bui et al., 2023; Cela et al., 2022; Diaz-Martínez et al., 2021; Lozano et al., 2023; Moreno et al., 2022; Muraoka et al., 2023; Takimoto et al., 2023), little progress has been made to further our knowledge regarding the specific roles that dysbiotic or pathogenic taxa may play in reproductive outcomes. Here, we applied metagenomics and transcriptomics to endometrial samples of 76 patients undergoing IVF treatment to explore the bacterial communities associated with reproductive failure and highlight the central bacterial functions at play. In brief, we described endometrial microbiome CSTs for the first time and reported a tendency for higher chance of reproductive failure among patients who bear communities enriched in *L. iners* or *G. vaginalis*. We also revealed differential functional networks present in endometrial samples, reporting a distinct genetic contribution from bacteria associated with reproductive failure or reproductive success. Lastly, host transcriptomics data suggested alteration in immune pathways and GPCR signaling may be the most affected by dysbiosis in patients with reproductive failure.

Studies regarding bacterial community composition and structure from distinct parts of the female genital tract have reported differences between the vaginal and the endometrial microbiota (C. Chen et al., 2017; Mitchell et al., 2015; Moreno et al., 2016). Here, we observed similar CSTs in endometrial samples to those found in vaginal samples. Pioneer studies have observed five different compositions of the vaginal microbiome - four dominated by *Lactobacillus* spp. (CST-I, *L. crispatus*; CST-II, *L. gasseri*; CST-III, *L. iners*; and CST-V, *L. jensenii*) and a fifth heterogeneous community

(CST-IV) displaying a higher abundance of strictly anaerobic bacteria such as *Atopobium*, *Prevotella*, *Streptococcus*, *Gardnerella*, and *Sneathia*, alongside a substantial decrease in *Lactobacillus* spp. (Gajer et al., 2012; Ravel et al., 2011). Nonetheless, additional assessment revealed the vaginal microbiota could be further divided into thirteen subgroups, depending on the relative abundance of the focal species but also considering minor abundance of other species (France et al., 2020). Compared to vaginal microbiotas, our hereby described CSTs-I and -III (dominated by *L. crispatus* and *L. iners*, respectively) displayed similar composition to those found in the vagina; nonetheless, our reported CST-II was composed of a highly homogeneous mix of several *Lactobacillus* species and other commensals. Additionally, four other CSTs were not dominated by *Lactobacillus* spp.: CST-IV (*E. coli*), CST-V (*G. vaginalis*), CST-VI (*L. lactis*), and CST-VII (*K. aerogenes*). Although these may be clear differences between vaginal and endometrial communities, it is important to note that technical diversity in experimental design and heterogeneous patient cohorts may have limited the ability to compare studies and allow definitive conclusions regarding endometrial microbiota (Toson et al., 2022).

Several studies have linked the presence of *L. crispatus*, the dominant taxa in CST-I, to a healthy status in reproductive-age women based on the bacteria's ability to prevent sexually transmitted infections and bacterial vaginosis, as well as an association with lower incidence of preterm birth (Fettweis et al., 2019; Golob et al., 2023; Ravel et al., 2021). In fact, overall abundance of *Lactobacillus* spp. has been proposed as a marker for reproductive success in both vaginal and endometrial microbiomes (Moreno et al., 2022; Petrova et al., 2015; Zeng et al., 2022), which would also benefit patients bearing an endometrial microbiota such as CST-II, characterized by a highly homogeneous presence of several *Lactobacilli* alongside other

commensals. Unlike other species from the genus, the role of *L. iners* in the FGT remains controversial, given the description of this bacteria as a vaginal symbiont and an opportunistic pathogen (N. Zheng et al., 2021). Interestingly, it has been detected in the transition between *Lactobacillus* and non-*Lactobacillus*-dominated CSTs (Gajer et al., 2012), and its unusual metabolism compared to other *Lactobacilli* may help understand why it is able to grow under these transition states (N. Zheng et al., 2021). Additionally, presence of *L. iners* in the vaginal microbiota has been associated with BV and preterm birth (Golob et al., 2023; Kindinger et al., 2017), increased susceptibility to *Chlamydia trachomatis* infection (van Houdt et al., 2018), pathogenesis of vulvovaginal candidiasis (Novak et al., 2022; Sabbatini et al., 2021), and may even indicate poor response to chemoradiation in cervical cancer (Colbert et al., 2023). Thus, our results add further evidence for the currently debated hypothesis that a *L. iners*-dominated microbiota may not be associated with urogenital health.

Lastly, CST-V was mainly composed of *Gardnerella vaginalis*, one of the most common pathogens in the female reproductive tract that causes BV (Swidsinski et al., 2013) and CE (Ravel et al., 2021). It exerts virulence by mechanisms that include the formation of biofilms and intrinsic resistance to specific antibiotics (Yeoman et al., 2010). Of note, several studies have associated the presence of *G. vaginalis* in the endometrium with reproductive failure (Garcia-Grau et al., 2019; X. Liu et al., 2021; Lozano et al., 2023; Moreno et al., 2022) and adenomyosis (Kunaseh et al., 2022), which agrees to the herein reported results. Nonetheless, 1 out of 8 patients who displayed high abundance of *G. vaginalis* (exceeding 20%) had a live birth, demonstrating the sole abundance of *G. vaginalis* does not completely impair a healthy pregnancy as previously suggested (Diaz-Martínez et al., 2021). In fact, we found samples from patients suffering from reproductive failure (BP, CM, or NP)

distributed across all bacterial communities (CST-I, -II, -III, and -V), demonstrating the impact that factors beyond the microbiome, including the host response, may play in fertility.

When looking for the composition of whole communities, we generally observed two big nodes of coexisting bacteria: one enriched in *Lactobacillus* species, and another comprising pathogens and dysbiotic microbes from genera such as *Gardnerella*, *Atopobium*, *Bifidobacterium*, and *Olsenella*. Interestingly, *L. iners* was not consistently found in a single network, showing a distinctive pattern that adds further evidence for the hypothesis that it might be a double player, adapting to the general environment and being able to thrive in transition states between eubiotic and dysbiotic microbiotas and highlighting the relevance of the gene profile of its strains to adapt to the environment. Importantly, strains from the same species have been demonstrated to trigger distinct and sometimes even opposing responses from the host, inducing different immunophenotypes of innate lymphoid cells (Hajjo & Geva-Zatorsky, 2021).

Accordingly, a significant strain-level diversity in *Lactobacillus* species has been demonstrated in vaginal samples, which may contribute to different microbiota functionality and consequently how it interacts with the host (Tortelli et al., 2021). Thus, it is feasible to speculate that differences in *L. iners* at the strain level could trigger distinct responses from the immune system, which would help explain why this species may have dichotomous functions and interacts with distinct groups of microbes in the endometrium. Overall, our co-occurrence network results confirm that not only isolated organisms, but entire bacterial communities deserve consideration when assessing the impacts of the endometrial microbiome in uterine pathophysiology.

When we expanded our network analysis and included functional data, we observed a higher number of taxa-function associations in all three groups of

reproductive failure (NP, BP, and CM) compared to reproductive success (LB), suggesting infertility-related dysbiosis may trigger an imbalance in molecular signaling not only between microbes, but also with the host. As an example, the cohesive positive associations of transposases with *Lactobacillus* spp. in different reproductive outcomes suggests these genes play an important role in balancing the endometrial microbiota composition. They are mobile genetic elements thought to have an important role in bacterial genome plasticity and host adaptation, given their rapid dissemination within and between genomes that leads to mutations and rearrangements that may offer selective advantage to thrive in a given ecosystem (Aziz et al., 2010), and could be beneficial for *Lactobacilli* to establish a colonization in the endometrial environment. Furthermore, ABC transporters had negative correlations with commensals (e.g. *Cutibacterium acnes*, *Lactobacillus gasseri*, and *Lactobacillus hominis*) in BP, but positively associated with dysbiotic bacteria (e.g. *Bifidobacterium* spp. and *G. vaginalis*) in CM. Importantly, these transporters have a wide range of functions, such as uptake of micronutrients, export of cytotoxic compounds secreted by host cells, and promote antibiotic resistance through multidrug efflux (Locher, 2016; Orelle et al., 2019). Lastly, sensor histidine kinases (HK) also had a positive association with dysbiotic bacteria such as *G. vaginalis*, *Bifidobacterium* spp. and *Olsenella* sp. As part of the two-component system, these proteins are essential for adaptation to environmental stress and cell survival in bacteria, and HK inhibitors are currently being developed as novel potential antimicrobial drugs (H. Chen et al., 2022; Khorchid & Ikura, 2006). Thus, these transporters and the HKs might significantly impact the ability of bacterial communities to inhabit a given ecosystem, and their association with specific microbes in the endometrium should be looked into further detail.

Importantly, previous research using shotgun sequencing has demonstrated distinct functional pathways in vaginal samples from women with preterm birth, as stratification of data into three Gene Ontology classifications (cellular component, biological process, and molecular functions) revealed functional diversity between the risk groupings (Feehily et al., 2020). Although this data was obtained from the vaginal microbiome, some results were in accordance with our dataset, and alteration in processes such as iron ion homeostasis, amino acid biosynthesis, carbohydrate metabolism, and tRNA aminoacylation were suggested as a potential functional signature of preterm birth (Feehily et al., 2020). Moreover, research using vaginal samples demonstrated transcriptional activities of *L. crispatus*, *L. iners*, and *G. vaginalis* vary significantly according to the taxonomic composition of the communities in which they reside, as *L. iners* and *G. vaginalis* had decreased expression of cholesterol-dependent cytolysins when co-existing with *Lactobacillus* spp., whereas their expression increases in the presence of other facultative and obligate anaerobes (France et al., 2022). Thus, it is most probable that not all abundant genes in our samples are transcriptionally active, and our study would benefit from transcriptomics analyses to gather more robust results about the uterine microenvironment and its relation to microbial communities.

Moreover, several genes encoding for virulence factors were significantly enriched in patients suffering from reproductive failure. Among the various genes that displayed differential abundance between groups, those coding for exotoxins, immune modulators, and factors associated with adherence, biofilm formation, and effector delivery system are some of the most representative. Given the significant impact of virulence on infection initiation, maintenance, and disease progression (Casadevall & Pirofski, 2009), gaining insight into which factors are associated with fertility onset is

crucial for further understanding of the underlying mechanisms of human conception. For instance, exotoxins help bacterial spread by triggering a cytokine storm and damaging tissues (Cavaillon, 2018), while the effector delivery system might act directly upon host cells, injecting molecules to disrupt their functions (Green & Mecsas, 2016).

On this path, virulence factors that promote biofilm formation have already been associated with the etiology of bacterial vaginosis, as it is associated with the presence of polymicrobial biofilm in the vagina and endometrium of affected women (Jung et al., 2017; Swidsinski et al., 2013). As biofilms shield bacteria against antibiotics, they significantly diminish therapy efficiency and therefore could be tightly associated with recurrent BV. Importantly, we observed an increased abundance of genes that may impact biofilm formation in our subset of patients suffering from reproductive failure, such as those associated with exopolysaccharide biosynthesis, phenazine production, glycogen/starch metabolism, and surface proteins like pili and fimbriae. Additionally, analysis of bacterial functions revealed enrichment of other genes previously associated with biofilm formation, such as isocitrate dehydrogenase and LPXTG, in patients with RF (Khodaparast et al., 2016; Prasad et al., 2017; Zhao et al., 2021). Although the impact of bacterial biofilms in the uterine cavity is yet to be characterized, it is feasible to speculate that they could act as a physical barrier against embryo implantation, negatively impacting pregnancy establishment and reproductive outcomes (Swidsinski et al., 2013). On the other hand, it has been described that biofilm formation by early colonizers may not be the direct cause of BV, but the responsible to prime the uterine environment for other bacteria, known as second colonizers, to thrive and trigger a more robust immune response and contribute to clinical symptoms (Muzny et al., 2020; Swidsinski et al., 2013). Nonetheless,

commensal bacteria, including *Lactobacillus* spp., also bear the capacity to produce biofilm, which could promote colonization and longer permanence in the mucosa of the host and avoid colonization by pathogenic bacteria (Salas-Jara et al., 2016). Thus, careful functional assessment should be conducted to estimate which genes from which bacterial species are responsible driving infertility, otherwise targeted therapy could be impacting the eubiotic endometrial microenvironment.

Several genes associated with the effector delivery system were also enriched in patients from the reproductive failure group, mainly those from types III, VI, and VII (T3SS, T6SS, and T7SS). While T3SS and T6SS are commonly found in Gram-negative bacteria (Coulthurst, 2019; Wagner et al., 2018), type VII secretion system is mainly found in Gram-positive bacteria and mycobacteria (Schindler et al., 2023). Their overall function is to inject virulence factors directly into host cells, which may have various targets. Once inside, these effector proteins manipulate host cell processes to promote bacterial survival and proliferation, evade immune detection, promote tissue damage, and establish infection. Thus, the effector delivery system represents a critical aspect of pathogenesis, enabling pathogens to exert precise control over host cell functions and exploit host resources for their own benefit (Coulthurst, 2019; Schindler et al., 2023; Wagner et al., 2018); and their increased abundance among patients from the RF group suggest their cell host modulation might significantly impact reproductive outcomes.

Additionally, various genes associated with nutritional factors were also overrepresented among patients suffering from reproductive failure. Among them, genes associated with iron transport are some of the most common, alongside transportation of manganese and magnesium ions. For instance, iron is a vital cofactor that enables a wide range of key metabolic bacterial activities. Nonetheless, its limited

abundance in some environments may significantly impact bacterial survival, and high affinity iron uptake systems play a significant role in maintaining homeostasis in iron-dependent bacteria, such as *E. coli* and *Salmonella typhimurium* (Pi & Helmann, 2017). Similarly, the presence of manganese is also crucial for bacterial survival, as it may function as a cofactor for diverse enzymes involved in metabolism, translation, and signaling (Bosma et al., 2021). Not surprisingly, macrophages have been reported to remove Mn and other essential metals inside the phagosome to starve intracellular pathogens (Juttukonda & Skaar, 2015).

It is important to mention that different bacteria may require distinct levels of these ions, depending on their metabolism. Remarkably, lactobacilli accumulate little to no Fe^{2+} , and bacteria belonging to this genus exhibit distinct metabolism patterns. For instance, *Lactobacillus* spp. are known to have a manganese-centric physiology and can tolerate very high amounts of Mn^{2+} , while accumulation of Mn^{2+} in nonmanganese-centric organisms may lead to cytotoxicity and cell death (Bosma et al., 2021). Additionally, the core genome of *L. crispatus* contains eight more ABC transporter genes when compared to *L. iners*, one of which is a known iron transporter; thus, it has been hypothesized that *L. crispatus* may sequester iron more effectively (France et al., 2016; Macklaim et al., 2011). Furthermore, *G. vaginalis* may potentially use three different mechanisms to acquire iron from a number of sources, demonstrating its high adaptability to distinct conditions (Jarosik et al., 1998). Ultimately, our results demonstrate bacterial communities from RF samples display distinct nutrient uptake mechanisms when compared to those communities from the LB group, possibly facilitating their growth in disrupted uterine conditions.

Building on this knowledge, targeting virulence factors has been suggested as an innovative approach to overcome the growing threat of antibiotic resistance (Gadar &

McCarthy, 2023), as impairing their binding to host targets would significantly diminish infection success without the need for proper microbe death. In fact, NGAs could be used to disrupt several microbial pathways and features, such as biofilm formation, adhesion to host cells and tissues, toxin production and secretion, quorum sensing, effector delivery system, among others. Considering we observed an increase in bacterial virulence factors among women with reproductive failure, the application of NGAs in infertile women might help reestablish a healthy uterine microenvironment by impairing bacterial ligands from reaching their biological effect on host cells, ultimately improving clinical outcomes. Nonetheless, it is crucial not to underestimate bacteria's ability to develop mechanisms to overcome NGA activity. The use of such targeted therapies requires in-depth knowledge regarding the specific pathways and their mechanisms involved in infertility. Moreover, it is important to keep in mind that commensal bacteria from the healthy microbiota also rely on these factors to successfully colonize the FGT, and could also have their behavior disrupted by these therapies. Thus, the selection of NGAs that target interactions specific to pathogens would not only mitigate secondary effects but also optimize clinical outcomes.

We also found patients with poor reproductive outcomes that displayed higher abundance of bacterial genes related to antibiotic resistance through several distinct processes, mainly those associated with efflux pumps; and more specifically the ones belonging to the ABC superfamily, but also others from the MFS, and RND transporter superfamilies. Importantly, members of the ABC family collectively mediate resistance to a wider range of antibiotic classes than any other group of resistance proteins, and are widespread in pathogenic bacteria (Sharkey & O'Neill, 2018). By actively pumping antibiotics out of bacterial cells, efflux pumps enable bacteria to evade the lethal effects of antibiotics, contributing to the development of antibiotic resistance (Gaurav

et al., 2023). Nonetheless, due to their wide range of action, efflux pumps may also be involved in multiple processes that facilitate virulence and colonization establishment, such as biofilm formation, transport of toxins for tissue invasion, quorum sensing, and acquisition of essential nutrients such as vitamins and transition metals (Gaurav et al., 2023). As an example, the enterobactin secreted by RND transporters in *E. coli* and other bacteria plays an important role for infection because it facilitates bacterial survival in low iron environments (Colclough et al., 2020). Additionally, MFS pumps have been demonstrated to modulate host innate immunity, facilitating bacterial spread and tissue invasion (Crimmins et al., 2008). Thus, extensive research has focused on the development of efflux pump inhibitors to significantly advance the treatment of bacterial infections (Garmory & Titball, 2004; Gaurav et al., 2023). Considering the increasing amount of data regarding the influence of the endometrial microbiome on reproductive medicine and the current use of antibiotics during fertility treatment, these features might be of great importance to clinical practice. For instance, antibiotic treatment significantly improved pregnancy and live birth rates among women with chronic endometritis and unexplained cause of infertility (Cicinelli et al., 2018). Additionally, antibiotic and probiotic administration successfully improved pregnancy and live birth rates among women with RIF and an abnormal endometrial microbiota (Iwami et al., 2023). Thus, genes related to antibiotic resistance may represent a toll on infertile patients, impairing antibiotic treatment efficacy and delaying reproductive success.

Our metagenomics results are the first to shed light on the structure of whole communities from the endometrial microbiome and also on the actual impact of these microbes in infertility, building a database that could give rise to a molecular

fingerprint in which detection of specific bacterial genes could be used as biomarkers for poor reproductive outcomes.

It has been described that the microbiome may modulate several host mechanisms related to chromatin organization, DNA methylation, and protein expression to create a more beneficial niche for infection (Nichols & Davenport, 2021). Importantly, microbiome-induced changes in gene expression and protein structure have been associated with the pathophysiology of different kinds of cancer (M. Guo et al., 2023; Park et al., 2022), and recent research has also unveiled mechanisms of microbiome-related gynecological diseases, such as BV and endometriosis (Berard et al., 2023; Muraoka et al., 2023), demonstrating the great impact microbes may have on the host cell physiology. In turn, a host response is also triggered against pathogens, adding an extra layer of complexity in the two-way microbiome-host interaction. In our subset of patients, we have found several genes related to immune response were differentially expressed in patients who suffered from reproductive failure in comparison to those who had a live birth. Accordingly, immune signaling has been described to be tightly regulated in the FGT: although immune tolerance is important for maintaining commensal bacteria (Nutsch & Hsieh, 2012), alterations in the microbiota may modulate pro-inflammatory factors, leading to a disruption in epithelial barrier and high levels of inflammation (Shannon et al., 2017; J. Wang et al., 2021). Consistently, vaginal dysbiosis has been significantly associated with the expression of proinflammatory cytokines such as IL-1 β , IL-6 and MIP-1 β in women suffering from preterm birth (Fettweis et al., 2019). Also, an *in vitro* model of bacterial colonization demonstrated co-cultivation of *Lactobacillus rhamnosus* with endometrial epithelial cells significantly modulates IL-6, IL-8, and other interleukin levels, suggesting its feasible use as a probiotic in gynecological health (Chenoll et al., 2019).

Noteworthy, bacteria may also modulate immune signaling through outer membrane vesicles, which contain several components such as proteins, metabolites, and short interfering RNAs that can target specific host signaling pathways and facilitate chronic infection (Charpentier et al., 2023). Additionally, the immune system should be tightly regulated during pregnancy to maintain tolerance towards the growing fetus, while still keeping up the defenses against pathogens. Thus, alterations in immunity pathways may associate with increased risk to subsequent infections that could significantly impair reproductive outcomes, but also impair embryo implantation and pregnancy development (Chazara et al., 2011; Cicinelli et al., 2018; Mor & Cardenas, 2010; Ravel et al., 2021). Conversely, we found patients from the NP group to display a downregulation of the *CCL2* gene, which may be expressed in decidual tissue, the myometrium, and others; and alterations in its signaling have been reported to be associated with adverse pregnancy outcomes such as spontaneous abortion, preeclampsia and preterm labor (Lin et al., 2023). Additionally, we also reported several other chemokines, interleukins, and cytokines to be differentially expressed among groups (Tables 22-24). While there are no previous studies specifically linking the genes we identified to human fertility, other similar molecules (such as IL-4, IL-5, IL-10, TNF- α , CX3CL1, and CLL4) have been described to play an important role in promoting trophoblast migration, embryo implantation, and endometrial cell remodeling (Dimitriadis et al., 2005; Hannan et al., 2006; Pantos et al., 2022; Piccinni et al., 2021; Złotkowska & Andronowska, 2019). Our results provide further evidence that altered signaling in these compounds may be associated with poor reproductive outcomes. Thus, our general results agree with previous literature, suggesting a disruption in immune signaling may partially explain reproductive failure in some of our patients.

Moreover, our subset of patients who suffer from reproductive failure also displayed differential expression of multiple GPCR-coding genes. GPCRs comprise the largest family of transmembrane proteins in the human genome, with over 800 members that perform a wide range of functions, transducing signals from several ligands such as chemokines, fatty acids, and peptides (Lagerström & Schiöth, 2008). Importantly, these receptors also transduce signals derived from several sexual hormones like follicle-stimulating hormone, luteinizing hormone, human chorionic gonadotropin, and gonadotropin-releasing hormone (Lagerström & Schiöth, 2008). Not surprisingly, these receptors mediate therapeutic effects of around 30% of the marketed drugs nowadays, as treatment acting via GPCRs may be used in conditions such as type 2 diabetes, cancer, osteoporosis, and hypertension (Hauser et al., 2017). More recently, the role of GPCRs has also extended to the urogenital tract, as research shows its relation to endometrial cell survival, menstrual cycle, and decidualization (Aikawa et al., 2017; Sprague et al., 2021; Totorikaguena et al., 2017), but also to gynecological diseases such as preeclampsia, endometriosis, and adenomyosis (Aghajanova et al., 2016; Xu et al., 2021; Yin et al., 2022b). Moreover, drug targeting of GPCRs has also been proposed as a feasible treatment for these conditions (Conrad, 2016; Tapmeier et al., 2021), demonstrating a significant growth of interest in these proteins in gynecology research.

Conversely, some GPCRs found to be differentially expressed in our study group have also been reported in different contexts of reproductive medicine. For instance, a study assessing endometrial samples from 40 patients with RIF and 27 fertile women revealed a lower expression of angiotensin II receptor (*AGTR2*) in patients suffering from RIF when compared to the fertile group. These results suggest the local renin-angiotensin system might be dysregulated in the peri-implantation endometrium, and

could potentially lead to abnormal angiogenesis and poor implantation outcomes after embryo transfer in patients undergoing IVF treatment (Qi et al., 2023). Although we observed an upregulation in the BP group when compared to patients who achieved a LB, we hypothesize that altered microvessel density arising from an abnormal angiotensin signaling may be associated with reproductive failure and pregnancy outcomes. Secondly, we also observed a downregulation in formyl peptide receptor 1 (*FPR1*) expression among patients from the NP group. Interestingly, histological assessment of ectopic endometrium from 18 patients with endometriosis and eutopic endometrium from 10 women without endometriosis revealed a marked increase in *FPR1* among epithelial cells from the endometriosis group, which indicates an inflammatory disorder in these patients (Paula et al., 2015). Thus, our data agrees with previous literature and suggests disruption in the annexin A1-FPR1 system may play a significant role in human fertility. We also reported an increased expression of *GLP2R* among patients suffering from clinical miscarriage. Interestingly, glucagon signaling has been associated with anti-inflammatory and anti-fibrotic effects in the gonads and the endometrium affected by obesity, diabetes, and polycystic ovary syndrome, and its inhibitors have significant potential to reverse diabetes-related fertility problems and polycystic ovary morphology (Jensterle et al., 2019). It is important to note that bacteria may mimic human ligands and trigger/impair GPCR signaling (Cohen et al., 2017). Thus, host genetic background and basal expression of GPCRs might also increase host sensibility to certain strains, which could help explain why some patients bearing the same bacterial communities have distinct reproductive outcomes.

Taking all of these into consideration, disrupted GPCR and immune signaling in our subset of patients agrees with previous literature, providing insights into the

pathophysiology of infertility and adding further evidence for the development of novel targeted therapies. Nonetheless, since the inhibition of human receptors may lead to various side effects, targeting bacterial metabolites known to alter host signaling, instead of the receptors themselves, represents a feasible and safer approach to prevent bacterial ligands from actively modulating cell host signaling pathways.

Most published articles focusing on microbiota employ 16S rRNA gene sequencing; however, our study assessed bacterial community composition using whole metagenomic sequencing, which not only enables the assessment of CSTs through the detection of bacteria at the species level but also allows comprehension of whole microbial genomes present in a set of samples. To the best of our knowledge, our study is the first to not rely solely on 16S rRNA gene sequencing to assess the impact of endometrial microbiota composition on reproductive outcomes. Further analyses using WMS allowed the association of several bacterial genes with reproductive failure, which also remained unreported. Additional strengths of our work included (i) personalized embryo transfer and collection of samples at the time of maximum endometrial receptivity during hormone replacement therapy, which mitigate further discrepancies in transcriptomic profile due to the time of sampling among patients; and (ii) the use of blanks and a big reading depth for the whole metagenome sequencing, which enabled the analysis of a low biomass microbiome such as the one found within the uterine cavity.

The main limitations of our study included the small number of samples in the BP and CM groups, which impaired the assessment of significant differences in microbiota composition and function among samples associated with reproductive failure. Nonetheless, it should be noted that the abundance of genes is not always indicative of the microbiome's transcriptional activity, and transcriptomic data would be a greater

tool to assess the ecology of endometrial microbiota. A continuum of microbiota has been proposed for the female genital tract; however, the endometrium represents a low biomass body site, and by assessing the endometrial microbiota through transcervical collection contamination cannot be fully discarded (Toson et al., 2022). However, this approach represents the only means to assess the endometrial microbiota in women undergoing ART, and clinical sampling was conducted cautiously following standardized protocols to prevent cross-contamination. Noteworthy, not all transferred embryos had been genetically assessed, meaning embryo ploidy might be a confounding factor for the impact of microbiota composition in reproductive outcomes, and may explain high reproductive failure rates among all the patients.

Regarding functional analyses of the microbiome, in our study we used Diamond/BLAST pipelines which is a very sensitive method for microbiome studies; however, it lacked power to assign bacterial genes to the taxa they come from, so we were only able to imply the taxa-function through correlation networks. Additionally, the very small number of bacterial reads in MTX analyses was not sufficient for full coverage of bacterial genomes, impairing a thorough assessment of microbial gene expression. Additionally, although metatranscriptomics results provide valuable insights into disrupted pathways in endometrial cells, integration analyses should be performed to fully elucidate which of these changes are triggered by bacteria or microbe-derived metabolites. Future research should be conducted to assess these topics more in detail. Moving forward, single-cell approaches targeting the human microbiome and endometrial cells should gather robust information regarding the molecular mechanisms underlying infertility. Additionally, *in vitro* models of endometrial infection coupled to gene editing should shed light on the specific pathways that have a greater impact on the pathophysiology of infertility.

Our findings of differential microbe genetic composition and expression of host genes in endometrial samples from patients with different reproductive outcomes may serve as a starting point for future research, and a schematic overview of all findings can be seen in Figure 47. Further assessment should shed light on new biomarkers that could give rise to diagnostic methods and targeted therapies to prevent microbiome dysbiosis from impairing human fertility, improving clinical outcomes when combined with standardized protocols.

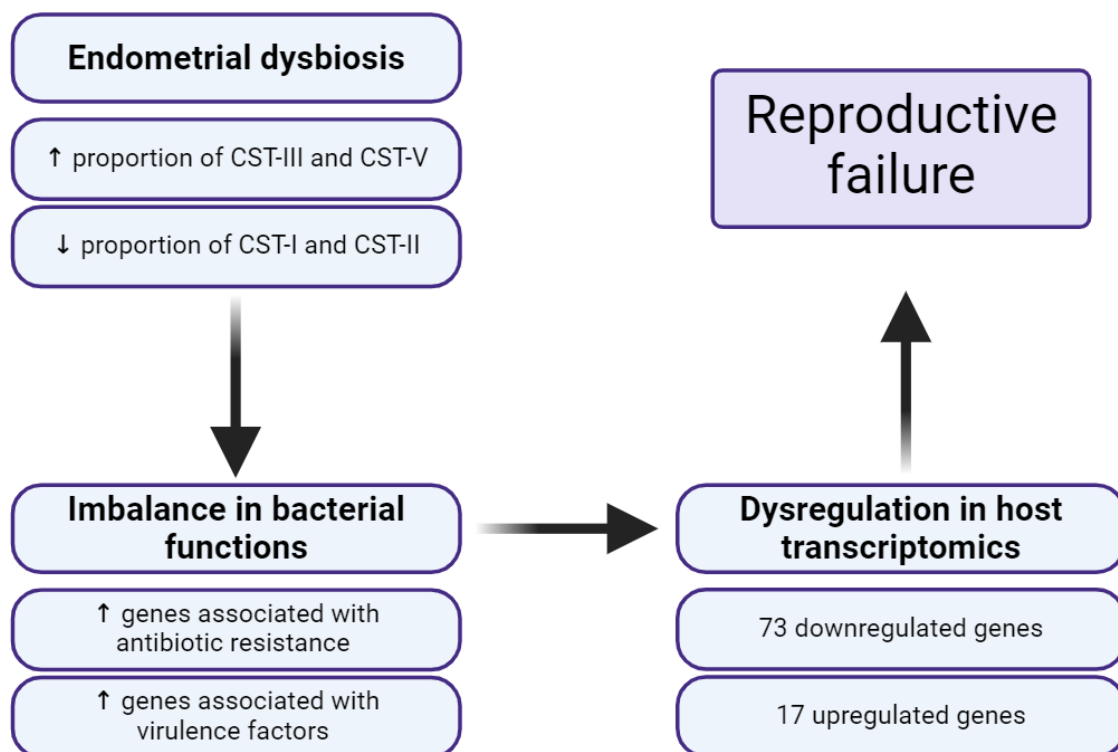


Figure 47. Overview of the most important findings from women who suffer from reproductive failure.

7. CONCLUSIONS

- The taxonomic study results using WMS confirm the existence of a differential microbiota at the species level in patients with reproductive failure, characterized by a significant enrichment of *G. vaginalis*, *B. breve*, *B. bifidum*, and *A. vaginae*, among other dysbiotic bacteria.
- CST composition in endometrial samples reveals four major communities, mainly inhabited by *L. crispatus*, *L. iners*, *G. vaginalis*, or a complex and homogeneous mixture of commensals. Communities dominated by *G. vaginalis* or *L. iners* tend to be associated with poorer reproductive outcomes when compared to those dominated by *L. crispatus* and other commensal communities.
- Functional WMS studies highlight the importance of considering the entire bacterial community to understand its impact on reproductive health, as the same bacterium can be associated positively (coexistence) or negatively (exclusion) with other bacteria, leading to differential genomic profiles in various reproductive outcomes.
- Antibiotic resistance genes are significantly enriched in the microbiome of patients suffering from reproductive failure, particularly those associated with antibiotic efflux pumps mediated by ABC, MFS and RND transporters.
- Virulence factors are significantly enriched in bacteria present in patients with reproductive failure. Most of these factors are related to immune modulation,

adherence and biofilm formation, nutritional factors, and systems for the delivery of metabolites and molecules.

- In the endometrium of patients with reproductive failure, there is a broad dysregulation of genes involved in the immune system, including humoral and cellular immunity, cytokine and chemokine secretion, phagocytosis, and bacterial defense response.
- Women suffering from reproductive failure also exhibit altered expression of several GPCRs in the endometrium, such as receptors for free fatty acids, angiotensin, glucagon, and hydroxycarboxylic acid.

8. REFERENCES

- Abd El Aziz, M. A., Sharifipour, F., Abedi, P., Jahanfar, S., & Judge, H. M. (2019). Secnidazole for treatment of bacterial vaginosis: a systematic review. *BMC Women's Health*, 19(1), 121. <https://doi.org/10.1186/s12905-019-0822-2>
- Abdallah, M., Benoliel, C., Drider, D., Dhulster, P., & Chihib, N.-E. (2014). Biofilm formation and persistence on abiotic surfaces in the context of food and medical environments. *Archives of Microbiology*, 196(7), 453–472. <https://doi.org/10.1007/s00203-014-0983-1>
- Abisado, R. G., Benomar, S., Klaus, J. R., Dandekar, A. A., & Chandler, J. R. (2018). Bacterial Quorum Sensing and Microbial Community Interactions. *MBio*, 9(3), e02331. <https://doi.org/10.1128/mBio.02331-17>
- Achilles, S. L., Austin, M. N., Meyn, L. A., Mhlanga, F., Chirenje, Z. M., & Hillier, S. L. (2018). Impact of contraceptive initiation on vaginal microbiota. *American Journal of Obstetrics and Gynecology*, 218(6), 622.e1-622.e10. <https://doi.org/10.1016/j.ajog.2018.02.017>
- Aghajanova, L., Altmäe, S., Kasvandik, S., Salumets, A., Stavreus-Evers, A., & Giudice, L. C. (2016). Stanniocalcin-1 expression in normal human endometrium and dysregulation in endometriosis. *Fertility and Sterility*, 106(3), 681-691.e1. <https://doi.org/10.1016/j.fertnstert.2016.05.023>
- Agus, A., Clément, K., & Sokol, H. (2021). Gut microbiota-derived metabolites as central regulators in metabolic disorders. *Gut*, 70(6), 1174–1182. <https://doi.org/10.1136/gutjnl-2020-323071>

- Aikawa, S., Kano, K., Inoue, A., Wang, J., Saigusa, D., Nagamatsu, T., Hirota, Y., Fujii, T., Tsuchiya, S., Taketomi, Y., Sugimoto, Y., Murakami, M., Arita, M., Kurano, M., Ikeda, H., Yatomi, Y., Chun, J., & Aoki, J. (2017). Autotaxin–lysophosphatidic acid– LPA_3 signaling at the embryo-epithelial boundary controls decidualization pathways. *The EMBO Journal*, 36(14), 2146–2160. <https://doi.org/10.15252/emboj.201696290>
- Alcock, B. P., Huynh, W., Chalil, R., Smith, K. W., Raphenya, A. R., Wlodarski, M. A., Edalatmand, A., Petkau, A., Syed, S. A., Tsang, K. K., Baker, S. J. C., Dave, M., McCarthy, M. C., Mukiri, K. M., Nasir, J. A., Golbon, B., Imtiaz, H., Jiang, X., Kaur, K., ... McArthur, A. G. (2023). CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Research*, 51(D1), D690–D699. <https://doi.org/10.1093/nar/gkac920>
- Aleti, G., Troyer, E. A., & Hong, S. (2023). G protein-coupled receptors: A target for microbial metabolites and a mechanistic link to microbiome-immune-brain interactions. *Brain, Behavior, & Immunity - Health*, 32, 100671. <https://doi.org/10.1016/j.bbih.2023.100671>
- Aziz, R. K., Breitbart, M., & Edwards, R. A. (2010). Transposases are the most abundant, most ubiquitous genes in nature. *Nucleic Acids Research*, 38(13), 4207–4217. <https://doi.org/10.1093/nar/gkq140>
- Baker, J. M., Chase, D. M., & Herbst-Kralovetz, M. M. (2018). Uterine Microbiota: Residents, Tourists, or Invaders? *Frontiers in Immunology*, 9(208). <https://doi.org/10.3389/fimmu.2018.00208>

- Baquero, F. (2021). Threats of antibiotic resistance: an obliged reappraisal. *International Microbiology*, 24(4), 499–506. <https://doi.org/10.1007/s10123-021-00184-y>
- Barton, H. A., Taylor, N. M., Lubbers, B. R., & Pemberton, A. C. (2006). DNA extraction from low-biomass carbonate rock: An improved method with reduced contamination and the low-biomass contaminant database. *Journal of Microbiological Methods*, 66(1), 21–31. <https://doi.org/10.1016/j.mimet.2005.10.005>
- Benner, M., Ferwerda, G., Joosten, I., & van der Molen, R. G. (2018). How uterine microbiota might be responsible for a receptive, fertile endometrium. *Human Reproduction Update*, 24(4). <https://doi.org/10.1093/humupd/dmy012>
- Berard, A. R., Brubaker, D. K., Birse, K., Lamont, A., Mackelprang, R. D., Noël-Romas, L., Perner, M., Hou, X., Irungu, E., Mugo, N., Knodel, S., Muwonge, T. R., Katabira, E., Hughes, S. M., Levy, C., Calienes, F. L., Lauffenburger, D. A., Baeten, J. M., Celum, C., ... Burgener, A. D. (2023). Vaginal epithelial dysfunction is mediated by the microbiome, metabolome, and mTOR signaling. *Cell Reports*, 42(5), 112474. <https://doi.org/10.1016/j.celrep.2023.112474>
- Biazzo, M., & Deidda, G. (2022). Fecal Microbiota Transplantation as New Therapeutic Avenue for Human Diseases. *Journal of Clinical Medicine*, 11(14), 4119. <https://doi.org/10.3390/jcm11144119>
- Bosma, E. F., Rau, M. H., van Gijtenbeek, L. A., & Siedler, S. (2021). Regulation and distinct physiological roles of manganese in bacteria. *FEMS Microbiology Reviews*, 45(6). <https://doi.org/10.1093/femsre/fuab028>

- Bourdon, M., Santulli, P., Jeljeli, M., Vannuccini, S., Marcellin, L., Doridot, L., Petraglia, F., Batteux, F., & Chapron, C. (2021). Immunological changes associated with adenomyosis: a systematic review. *Human Reproduction Update*, 27(1), 108–129. <https://doi.org/10.1093/humupd/dmaa038>
- Bourdon, M., Santulli, P., Marcellin, L., Maignien, C., Maitrot-Mantelet, L., Bordonne, C., Plu Bureau, G., & Chapron, C. (2021). Adenomyosis: An update regarding its diagnosis and clinical features. *Journal of Gynecology Obstetrics and Human Reproduction*, 50(10), 102228. <https://doi.org/10.1016/j.jogoh.2021.102228>
- Brooks, J. P., Edwards, D. J., Blithe, D. L., Fettweis, J. M., Serrano, M. G., Sheth, N. U., Strauss, J. F., Buck, G. A., & Jefferson, K. K. (2017). Effects of combined oral contraceptives, depot medroxyprogesterone acetate and the levonorgestrel-releasing intrauterine system on the vaginal microbiome. *Contraception*, 95(4), 405–413. <https://doi.org/10.1016/j.contraception.2016.11.006>
- Buchfink, B., Reuter, K., & Drost, H.-G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nature Methods*, 18(4), 366–368. <https://doi.org/10.1038/s41592-021-01101-x>
- Bui, B. N., van Hoogenhuijze, N., Viveen, M., Mol, F., Teklenburg, G., de Bruin, J.-P., Besselink, D., Brentjens, L. S., Mackens, S., Rogers, M. R. C., Steba, G. S., Broekmans, F., Paganelli, F. L., & van de Wijgert, J. H. H. M. (2023). The endometrial microbiota of women with or without a live birth within 12 months after a first failed IVF/ICSI cycle. *Scientific Reports*, 13(1), 3444. <https://doi.org/10.1038/s41598-023-30591-2>

- Burgueño, J. F., & Abreu, M. T. (2020). Epithelial Toll-like receptors and their role in gut homeostasis and disease. *Nature Reviews Gastroenterology & Hepatology*, 17(5), 263–278. <https://doi.org/10.1038/s41575-019-0261-4>
- Carosso, A., Revelli, A., Gennarelli, G., Canosa, S., Cosma, S., Borella, F., Tancredi, A., Pascherio, C., Boatti, L., Zanotto, E., Sidoti, F., Bottino, P., Costa, C., Cavallo, R., & Benedetto, C. (2020). Controlled ovarian stimulation and progesterone supplementation affect vaginal and endometrial microbiota in IVF cycles: a pilot study. *Journal of Assisted Reproduction and Genetics*, 37(9), 2315–2326. <https://doi.org/10.1007/s10815-020-01878-4>
- Casadevall, A., & Pirofski, L. (2009). Virulence factors and their mechanisms of action: the view from a damage–response framework. *Journal of Water and Health*, 7(S1), S2–S18. <https://doi.org/10.2166/wh.2009.036>
- Cavaillon, J.-M. (2018). Exotoxins and endotoxins: Inducers of inflammatory cytokines. *Toxicon*, 149, 45–53. <https://doi.org/10.1016/j.toxicon.2017.10.016>
- Cela, V., Daniele, S., Obino, M. E. R., Ruggiero, M., Zappelli, E., Ceccarelli, L., Papini, F., Marzi, I., Scarfò, G., Tosi, F., Franzoni, F., Martini, C., & Artini, P. G. (2022). Endometrial Dysbiosis Is Related to Inflammatory Factors in Women with Repeated Implantation Failure: A Pilot Study. *Journal of Clinical Medicine*, 11(9), 2481. <https://doi.org/10.3390/jcm11092481>
- Chakravorty, S., Helb, D., Burday, M., Connell, N., & Alland, D. (2007). A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *Journal of Microbiological Methods*, 69(2), 330–339. <https://doi.org/10.1016/j.mimet.2007.02.005>

- Charpentier, L. A., Dolben, E. F., Hendricks, M. R., Hogan, D. A., Bomberger, J. M., & Stanton, B. A. (2023). Bacterial Outer Membrane Vesicles and Immune Modulation of the Host. *Membranes*, 13(9), 752. <https://doi.org/10.3390/membranes13090752>
- Chazara, O., Xiong, S., & Moffett, A. (2011). Maternal KIR and fetal HLA-C: a fine balance. *Journal of Leukocyte Biology*, 90(4), 703–716. <https://doi.org/10.1189/jlb.0511227>
- Chee, W. J. Y., Chew, S. Y., & Than, L. T. L. (2020). Vaginal microbiota and the potential of Lactobacillus derivatives in maintaining vaginal health. *Microbial Cell Factories*, 19(1), 203. <https://doi.org/10.1186/s12934-020-01464-4>
- Chen, C., Song, X., Wei, W., Zhong, H., Dai, J., Lan, Z., Li, F., Yu, X., Feng, Q., Wang, Z., Xie, H., Chen, X., Zeng, C., Wen, B., Zeng, L., Du, H., Tang, H., Xu, C., Xia, Y., ... Jia, H. (2017). The microbiota continuum along the female reproductive tract and its relation to uterine-related diseases. *Nature Communications*, 8(1), 875. <https://doi.org/10.1038/s41467-017-00901-0>
- Chen, H., Rosen, C. E., González-Hernández, J. A., Song, D., Potempa, J., Ring, A. M., & Palm, N. W. (2023). Highly multiplexed bioactivity screening reveals human and microbiota metabolome-GPCRs interactions. *Cell*, 186(14), 3095-3110.e19. <https://doi.org/10.1016/j.cell.2023.05.024>
- Chen, H., Yu, C., Wu, H., Li, G., Li, C., Hong, W., Yang, X., Wang, H., & You, X. (2022). Recent Advances in Histidine Kinase-Targeted Antimicrobial Agents. *Frontiers in Chemistry*, 10. <https://doi.org/10.3389/fchem.2022.866392>

- Chen, L. (2005). VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Research*, 33(Database issue), D325–D328. <https://doi.org/10.1093/nar/gki008>
- Chen, T., Xia, C., Hu, H., Wang, H., Tan, B., Tian, P., Zhao, X., Wang, L., Han, Y., Deng, K.-Y., Wei, H., & Xin, H.-B. (2021). Dysbiosis of the rat vagina is efficiently rescued by vaginal microbiota transplantation or probiotic combination. *International Journal of Antimicrobial Agents*, 57(3), 106277. <https://doi.org/10.1016/j.ijantimicag.2021.106277>
- Cheng, X., Huang, Z., Xiao, Z., & Bai, Y. (2022). Does antibiotic therapy for chronic endometritis improve clinical outcomes of patients with recurrent implantation failure in subsequent IVF cycles? A systematic review and meta-analysis. *Journal of Assisted Reproduction and Genetics*, 39(8), 1797–1813. <https://doi.org/10.1007/s10815-022-02558-1>
- Chenoll, E., Moreno, I., Sánchez, M., Garcia-Grau, I., Silva, Á., González-Monfort, M., Genovés, S., Vilella, F., Seco-Durban, C., Simón, C., & Ramón, D. (2019). Selection of New Probiotics for Endometrial Health. *Frontiers in Cellular and Infection Microbiology*, 9, 114. <https://doi.org/10.3389/fcimb.2019.00114>
- Cicinelli, E., Matteo, M., Tinelli, R., Lepera, A., Alfonso, R., Indraccolo, U., Marrocchella, S., Greco, P., & Resta, L. (2015). Prevalence of chronic endometritis in repeated unexplained implantation failure and the IVF success rate after antibiotic therapy. *Human Reproduction*, 30(2), 323–330. <https://doi.org/10.1093/humrep/deu292>
- Cicinelli, E., Matteo, M., Tinelli, R., Pinto, V., Marinaccio, M., Indraccolo, U., De Ziegler, D., & Resta, L. (2014). Chronic Endometritis Due to Common Bacteria Is

- Prevalent in Women With Recurrent Miscarriage as Confirmed by Improved Pregnancy Outcome After Antibiotic Treatment. *Reproductive Sciences*, 21(5). <https://doi.org/10.1177/1933719113508817>
- Cicinelli, E., Matteo, M., Trojano, G., Mitola, P. C., Tinelli, R., Vitagliano, A., Crupano, F. M., Lepera, A., Miragliotta, G., & Resta, L. (2018). Chronic endometritis in patients with unexplained infertility: Prevalence and effects of antibiotic treatment on spontaneous conception. *American Journal of Reproductive Immunology*, 79(1). <https://doi.org/10.1111/aji.12782>
- Clarridge, J. E. (2004). Impact of 16S rRNA Gene Sequence Analysis for Identification of Bacteria on Clinical Microbiology and Infectious Diseases. *Clinical Microbiology Reviews*, 17(4), 840–862. <https://doi.org/10.1128/CMR.17.4.840-862.2004>
- Co, E. C., Gormley, M., Kapidzic, M., Rosen, D. B., Scott, M. A., Stolp, H. A. R., McMaster, M., Lanier, L. L., Barcena, A., & Fisher, S. J. (2013). Maternal Decidual Macrophages Inhibit NK Cell Killing of Invasive Cytotrophoblasts During Human Pregnancy. *Biology of Reproduction*, 88(6), 155–155. <https://doi.org/10.1095/biolreprod.112.099465>
- Cohen, L. J., Esterhazy, D., Kim, S.-H., Lemetre, C., Aguilar, R. R., Gordon, E. A., Pickard, A. J., Cross, J. R., Emiliano, A. B., Han, S. M., Chu, J., Vila-Farres, X., Kaplitt, J., Rogoz, A., Calle, P. Y., Hunter, C., Bitok, J. K., & Brady, S. F. (2017). Commensal bacteria make GPCR ligands that mimic human signalling molecules. *Nature*, 549(7670), 48–53. <https://doi.org/10.1038/nature23874>
- Colbert, L. E., El Alam, M. B., Wang, R., Karpinets, T., Lo, D., Lynn, E. J., Harris, T. A., Elnaggar, J. H., Yoshida-Court, K., Tomasic, K., Bronk, J. K., Sammouri, J., Yanamandra, A. V., Olvera, A. V., Carlin, L. G., Sims, T., Delgado Medrano, A. Y.,

- Nappravnik, T. C., O'Hara, M., ... Klopp, A. H. (2023). Tumor-resident *Lactobacillus* iners confer chemoradiation resistance through lactate-induced metabolic rewiring. *Cancer Cell*, 41(11), 1945-1962.e11. <https://doi.org/10.1016/j.ccell.2023.09.012>
- Colclough, A. L., Alav, I., Whittle, E. E., Pugh, H. L., Darby, E. M., Legood, S. W., McNeil, H. E., & Blair, J. M. (2020). RND efflux pumps in Gram-negative bacteria; regulation, structure and role in antibiotic resistance. *Future Microbiology*, 15(2), 143–157. <https://doi.org/10.2217/fmb-2019-0235>
- Colosimo, D. A., Kohn, J. A., Luo, P. M., Piscotta, F. J., Han, S. M., Pickard, A. J., Rao, A., Cross, J. R., Cohen, L. J., & Brady, S. F. (2019). Mapping Interactions of Microbial Metabolites with Human G-Protein-Coupled Receptors. *Cell Host & Microbe*, 26(2), 273-282.e7. <https://doi.org/10.1016/j.chom.2019.07.002>
- Conrad, K. P. (2016). G-Protein-coupled receptors as potential drug candidates in preeclampsia: targeting the relaxin/insulin-like family peptide receptor 1 for treatment and prevention. *Human Reproduction Update*, 22(5), 647–664. <https://doi.org/10.1093/humupd/dmw021>
- Corbett, G., Crosby, D., & McAuliffe, F. (2021). Probiotic therapy in couples with infertility: A systematic review. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 256, 95–100. <https://doi.org/10.1016/j.ejogrb.2020.10.054>
- Coulthurst, S. (2019). The Type VI secretion system: a versatile bacterial weapon. *Microbiology*, 165(5), 503–515. <https://doi.org/10.1099/mic.0.000789>
- Crimmins, G. T., Herskovits, A. A., Rehder, K., Sivick, K. E., Lauer, P., Dubensky, T. W., & Portnoy, D. A. (2008). *Listeria monocytogenes* multidrug resistance transporters activate a cytosolic surveillance pathway of innate immunity.

- Proceedings of the National Academy of Sciences*, 105(29), 10191–10196.
<https://doi.org/10.1073/pnas.0804170105>
- Critchley, H. O. D., Maybin, J. A., Armstrong, G. M., & Williams, A. R. W. (2020). Physiology of the Endometrium and Regulation of Menstruation. *Physiological Reviews*, 100(3), 1149–1179. <https://doi.org/10.1152/physrev.00031.2019>
- Cryan, J. F., O’Riordan, K. J., Cowan, C. S. M., Sandhu, K. V., Bastiaanssen, T. F. S., Boehme, M., Codagnone, M. G., Cusotto, S., Fulling, C., Golubeva, A. V., Guzzetta, K. E., Jaggar, M., Long-Smith, C. M., Lyte, J. M., Martin, J. A., Molinero-Perez, A., Moloney, G., Morelli, E., Morillas, E., ... Dinan, T. G. (2019). The Microbiota-Gut-Brain Axis. *Physiological Reviews*, 99(4), 1877–2013. <https://doi.org/10.1152/physrev.00018.2018>
- Csardi, G., & Nepusz, T. (2006). The Igraph Software Package for Complex Network Research. *InterJournal, Complex Systems*, 1695.
- Dalby, J., & Stoner, B. P. (2022). Sexually Transmitted Infections: Updates From the 2021 CDC Guidelines. *American Family Physician*, 105(5), 514–520.
- Deer, E., Herrock, O., Campbell, N., Cornelius, D., Fitzgerald, S., Amaral, L. M., & LaMarca, B. (2023). The role of immune cells and mediators in preeclampsia. *Nature Reviews Nephrology*, 19(4), 257–270. <https://doi.org/10.1038/s41581-022-00670-0>
- Díaz-Martínez, M. del C., Bernabeu, A., Lledó, B., Carratalá-Munuera, C., Quesada, J. A., Lozano, F. M., Ruiz, V., Morales, R., Llácer, J., Ten, J., Castillo, J. C., Rodríguez, A., Nouni-García, R., López-Pineda, A., Moliner, B., & Bernabeu, R. (2021). Impact of the Vaginal and Endometrial Microbiome Pattern on Assisted

- Reproduction Outcomes. *Journal of Clinical Medicine* 2021, 10(18), 4063.
<https://doi.org/10.3390/JCM10184063>
- Dimitriadis, E., White, C. A., Jones, R. L., & Salamonsen, L. A. (2005). Cytokines, chemokines and growth factors in endometrium related to implantation. *Human Reproduction Update*, 11(6), 613–630. <https://doi.org/10.1093/humupd/dmi023>
- Dorsam, R. T., & Gutkind, J. S. (2007). G-protein-coupled receptors and cancer. *Nature Reviews Cancer*, 7(2), 79–94. <https://doi.org/10.1038/nrc2069>
- Durand, G. A., Raoult, D., & Dubourg, G. (2019). Antibiotic discovery: history, methods and perspectives. *International Journal of Antimicrobial Agents*, 53(4), 371–382.
<https://doi.org/10.1016/j.ijantimicag.2018.11.010>
- Edgar, R. C. (2018). Updating the 97% identity threshold for 16S ribosomal RNA OTUs. *Bioinformatics*, 34(14), 2371–2375.
<https://doi.org/10.1093/bioinformatics/bty113>
- El-Sayed, A., Aleya, L., & Kamel, M. (2021a). Microbiota's role in health and diseases. *Environmental Science and Pollution Research*, 28(28), 36967–36983.
<https://doi.org/10.1007/s11356-021-14593-z>
- El-Sayed, A., Aleya, L., & Kamel, M. (2021b). The link among microbiota, epigenetics, and disease development. *Environmental Science and Pollution Research*, 28(23), 28926–28964. <https://doi.org/10.1007/s11356-021-13862-1>
- ESHRE Working Group on Recurrent Implantation Failure, Cimadomo, D., de los Santos, M. J., Griesinger, G., Lainas, G., Le Clef, N., McLernon, D. J., Montjean, D., Toth, B., Vermeulen, N., & Macklon, N. (2023). ESHRE good practice

- recommendations on recurrent implantation failure. *Human Reproduction Open*, 2023(3). <https://doi.org/10.1093/hropen/hoad023>
- Feehily, C., Crosby, D., Walsh, C. J., Lawton, E. M., Higgins, S., McAuliffe, F. M., & Cotter, P. D. (2020). Shotgun sequencing of the vaginal microbiome reveals both a species and functional potential signature of preterm birth. *Npj Biofilms and Microbiomes*, 6(1), 50. <https://doi.org/10.1038/s41522-020-00162-8>
- Fettweis, J. M., Serrano, M. G., Brooks, J. P., Edwards, D. J., Girerd, P. H., Parikh, H. I., Huang, B., Arodz, T. J., Edupuganti, L., Glascock, A. L., Xu, J., Jimenez, N. R., Vivadelli, S. C., Fong, S. S., Sheth, N. U., Jean, S., Lee, V., Bokhari, Y. A., Lara, A. M., ... Buck, G. A. (2019). The vaginal microbiome and preterm birth. *Nature Medicine*, 25(6), 1012–1021. <https://doi.org/10.1038/s41591-019-0450-2>
- Franasiak, J. M., Werner, M. D., Juneau, C. R., Tao, X., Landis, J., Zhan, Y., Treff, N. R., & Scott, R. T. (2016). Endometrial microbiome at the time of embryo transfer: next-generation sequencing of the 16S ribosomal subunit. *Journal of Assisted Reproduction and Genetics*, 33(1), 129–136. <https://doi.org/10.1007/s10815-015-0614-z>
- France, M. T., Fu, L., Rutt, L., Yang, H., Humphrys, M. S., Narina, S., Gajer, P. M., Ma, B., Forney, L. J., & Ravel, J. (2022). Insight into the ecology of vaginal bacteria through integrative analyses of metagenomic and metatranscriptomic data. *Genome Biology*, 23(1), 66. <https://doi.org/10.1186/s13059-022-02635-9>
- France, M. T., Ma, B., Gajer, P., Brown, S., Humphrys, M. S., Holm, J. B., Waetjen, L. E., Brotman, R. M., & Ravel, J. (2020). VALENCIA: a nearest centroid classification method for vaginal microbial communities based on composition. *Microbiome*, 8(1), 166. <https://doi.org/10.1186/s40168-020-00934-6>

- France, M. T., Mendes-Soares, H., & Forney, L. J. (2016). Genomic Comparisons of *Lactobacillus crispatus* and *Lactobacillus iners* Reveal Potential Ecological Drivers of Community Composition in the Vagina. *Applied and Environmental Microbiology*, 82(24), 7063–7073. <https://doi.org/10.1128/AEM.02385-16>
- Friedman, J., & Alm, E. J. (2012). Inferring Correlation Networks from Genomic Survey Data. *PLoS Computational Biology*, 8(9), e1002687. <https://doi.org/10.1371/journal.pcbi.1002687>
- Fujii, S., & Oguchi, T. (2023). Age- and endometrial microbiota-related delay in development of endometrial receptivity. *Reproductive Medicine and Biology*, 22(1). <https://doi.org/10.1002/rmb2.12523>
- Gadar, K., & McCarthy, R. R. (2023). Using next generation antimicrobials to target the mechanisms of infection. *Npj Antimicrobials and Resistance*, 1(1), 11. <https://doi.org/10.1038/s44259-023-00011-6>
- Gajer, P., Brotman, R. M., Bai, G., Sakamoto, J., Schütte, U. M. E., Zhong, X., Koenig, S. S. K., Fu, L., Ma, Z. (Sam), Zhou, X., Abdo, Z., Forney, L. J., & Ravel, J. (2012). Temporal Dynamics of the Human Vaginal Microbiota. *Science Translational Medicine*, 4(132), 132ra52. <https://doi.org/10.1126/scitranslmed.3003605>
- Galloway-Peña, J., & Hanson, B. (2020). Tools for Analysis of the Microbiome. *Digestive Diseases and Sciences*, 65(3), 674–685. <https://doi.org/10.1007/s10620-020-06091-y>
- Garcia-Grau, I., Perez-Villaroya, D., Bau, D., Gonzalez-Monfort, M., Vilella, F., Moreno, I., & Simon, C. (2019). Taxonomical and Functional Assessment of the Endometrial Microbiota in A Context of Recurrent Reproductive Failure: A Case Report. *Pathogens*, 8(4), 205. <https://doi.org/10.3390/pathogens8040205>

- Garmory, H. S., & Titball, R. W. (2004). ATP-Binding Cassette Transporters Are Targets for the Development of Antibacterial Vaccines and Therapies. *Infection and Immunity*, 72(12), 6757–6763. <https://doi.org/10.1128/IAI.72.12.6757-6763.2004>
- Garrido-Gómez, T., Castillo-Marco, N., Cordero, T., & Simón, C. (2022). Decidualization resistance in the origin of preeclampsia. *American Journal of Obstetrics and Gynecology*, 226(2), S886–S894. <https://doi.org/10.1016/j.ajog.2020.09.039>
- Garrido-Gomez, T., Dominguez, F., Lopez, J. A., Camafeita, E., Quiñonero, A., Martinez-Conejero, J. A., Pellicer, A., Conesa, A., & Simón, C. (2011). Modeling Human Endometrial Decidualization from the Interaction between Proteome and Secretome. *The Journal of Clinical Endocrinology & Metabolism*, 96(3), 706–716. <https://doi.org/10.1210/jc.2010-1825>
- Gaurav, A., Bakht, P., Saini, M., Pandey, S., & Pathania, R. (2023). Role of bacterial efflux pumps in antibiotic resistance, virulence, and strategies to discover novel efflux pump inhibitors. *Microbiology*, 169(5), 001333. <https://doi.org/10.1099/mic.0.001333>
- Gholiof, M., Adamson-De Luca, E., & Wessels, J. M. (2022). The female reproductive tract microbiotas, inflammation, and gynecological conditions. *Frontiers in Reproductive Health*, 4. <https://doi.org/10.3389/frph.2022.963752>
- Giudice, L. C., & Kao, L. C. (2004). Endometriosis. *The Lancet*, 364(9447), 1789–1799. [https://doi.org/10.1016/S0140-6736\(04\)17403-5](https://doi.org/10.1016/S0140-6736(04)17403-5)
- Glassing, A., Dowd, S. E., Galandiuk, S., Davis, B., & Chiodini, R. J. (2016). Inherent bacterial DNA contamination of extraction and sequencing reagents may affect

- interpretation of microbiota in low bacterial biomass samples. *Gut Pathogens*, 8, 24. <https://doi.org/10.1186/s13099-016-0103-7>
- Golob, J. L., Oskotsky, T. T., Tang, A. S., Roldan, A., Chung, V., Ha, C. W. Y., Wong, R. J., Flynn, K. J., Parraga-Leo, A., Wibrand, C., Minot, S. S., Oskotsky, B., Andreoletti, G., Kostj, I., Bletz, J., Nelson, A., Gao, J., Wei, Z., Chen, G., ... Sirota, M. (2023). Microbiome preterm birth DREAM challenge: Crowdsourcing machine learning approaches to advance preterm birth research. *Cell Reports Medicine*, 101350. <https://doi.org/10.1016/j.xcrm.2023.101350>
- Goswami, A. G., Basu, S., Banerjee, T., & Shukla, V. K. (2023). Biofilm and wound healing: from bench to bedside. *European Journal of Medical Research*, 28(1), 157. <https://doi.org/10.1186/s40001-023-01121-7>
- Grahn, N., Olofsson, M., Ellnebo-Svedlund, K., Monstein, H.-J., & Jonasson, J. (2003). Identification of mixed bacterial DNA contamination in broad-range PCR amplification of 16S rDNA V1 and V3 variable regions by pyrosequencing of cloned amplicons. *FEMS Microbiology Letters*, 219(1), 87–91. [https://doi.org/10.1016/S0378-1097\(02\)01190-4](https://doi.org/10.1016/S0378-1097(02)01190-4)
- Gras, E., Bergeron, E., Puges, M., Ducours, M., Leleux, C., Amoureux, L., Jean, B., Bendjelloul, I., Camelena, F., Chenouard, R., Mahieu, R., Lemenand, O., Toro, A., Lecoustumier, A., Lortholary, O., Rodriguez Nava, V., & Lebeaux, D. (2022). Identification of *Streptomyces* spp. in a Clinical Sample: Always Contamination? Results of a French Retrospective Study. *Open Forum Infectious Diseases*, 9(7). <https://doi.org/10.1093/ofid/ofac271>
- Graspeuntner, S., Loeper, N., Künzel, S., Baines, J. F., & Rupp, J. (2018). Selection of validated hypervariable regions is crucial in 16S-based microbiota studies of

- the female genital tract. *Scientific Reports*, 8(1), 9678.
<https://doi.org/10.1038/s41598-018-27757-8>
- Green, E. R., & Mecsas, J. (2016). Bacterial Secretion Systems: An Overview. *Microbiology Spectrum*, 4(1). <https://doi.org/10.1128/microbiolspec.VMBF-0012-2015>
- Guo, F., Si, C., Zhou, M., Wang, J., Zhang, D., Leung, P. C. K., Xu, B., & Zhang, A. (2018). Decreased PECAM1-mediated TGF- β 1 expression in the mid-secretory endometrium in women with recurrent implantation failure. *Human Reproduction*, 33(5), 832–843. <https://doi.org/10.1093/humrep/dey022>
- Guo, M., Yao, Z., Jiang, C., Songyang, Z., Gan, L., & Xiong, Y. (2023). Three-dimensional and single-cell sequencing of liver cancer reveals comprehensive host-virus interactions in HBV infection. *Frontiers in Immunology*, 14. <https://doi.org/10.3389/fimmu.2023.1161522>
- Häder, A., Schäuble, S., Gehlen, J., Thielemann, N., Buerfent, B. C., Schüller, V., Hess, T., Wolf, T., Schröder, J., Weber, M., Hünninger, K., Löffler, J., Vylkova, S., Panagiotou, G., Schumacher, J., & Kurzai, O. (2023). Pathogen-specific innate immune response patterns are distinctly affected by genetic diversity. *Nature Communications*, 14(1), 3239. <https://doi.org/10.1038/s41467-023-38994-5>
- Haggerty, C. L., Hillier, S. L., Bass, D. C., & Ness, R. B. (2004). Bacterial Vaginosis and Anaerobic Bacteria Are Associated with Endometritis. *Clinical Infectious Diseases*, 39(7), 990–995. <https://doi.org/10.1086/423963>
- Haggerty, C. L., Totten, P. A., Tang, G., Astete, S. G., Ferris, M. J., Norori, J., Bass, D. C., Martin, D. H., Taylor, B. D., & Ness, R. B. (2016). Identification of novel microbes associated with pelvic inflammatory disease and infertility. *Sexually*

- Transmitted Infections*, 92(6), 441–446. <https://doi.org/10.1136/sextrans-2015-052285>
- Hajjo, H., & Geva-Zatorsky, N. (2021). Strain-level immunomodulatory variation of gut bacteria. *FEBS Letters*, 595(9), 1322–1327. <https://doi.org/10.1002/1873-3468.14057>
- Hannan, N. J., Jones, R. L., White, C. A., & Salamonsen, L. A. (2006). The Chemokines, CX3CL1, CCL14, and CCL4, Promote Human Trophoblast Migration at the Feto-Maternal Interface¹. *Biology of Reproduction*, 74(5), 896–904. <https://doi.org/10.1095/biolreprod.105.045518>
- Harris-Tryon, T. A., & Grice, E. A. (2022). Microbiota and maintenance of skin barrier function. *Science*, 376(6596), 940–945. <https://doi.org/10.1126/science.abo0693>
- Hashimoto, T., & Kyono, K. (2019). Does dysbiotic endometrium affect blastocyst implantation in IVF patients? *Journal of Assisted Reproduction and Genetics*, 36(12), 2471–2479. <https://doi.org/10.1007/s10815-019-01630-7>
- Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schiöth, H. B., & Gloriam, D. E. (2017). Trends in GPCR drug discovery: new agents, targets and indications. *Nature Reviews Drug Discovery*, 16(12), 829–842. <https://doi.org/10.1038/nrd.2017.178>
- Hiergeist, A., Gläsner, J., Reischl, U., & Gessner, A. (2015). Analyses of Intestinal Microbiota: Culture versus Sequencing. *ILAR Journal*, 56(2), 228–240. <https://doi.org/10.1093/ilar/ilv017>
- Hirakawa, H., Kurushima, J., Hashimoto, Y., & Tomita, H. (2020). Progress Overview of Bacterial Two-Component Regulatory Systems as Potential Targets for

- Antimicrobial Chemotherapy. *Antibiotics*, 9(10), 635.
<https://doi.org/10.3390/antibiotics9100635>
- Hitch, T. C. A., Hall, L. J., Walsh, S. K., Leventhal, G. E., Slack, E., de Wouters, T., Walter, J., & Clavel, T. (2022). Microbiome-based interventions to modulate gut ecology and the immune system. *Mucosal Immunology*, 15(6), 1095–1113.
<https://doi.org/10.1038/s41385-022-00564-1>
- Hou, K., Wu, Z.-X., Chen, X.-Y., Wang, J.-Q., Zhang, D., Xiao, C., Zhu, D., Koya, J. B., Wei, L., Li, J., & Chen, Z.-S. (2022). Microbiota in health and diseases. *Signal Transduction and Targeted Therapy*, 7(1), 135. <https://doi.org/10.1038/s41392-022-00974-4>
- Hsieh, S. A., & Allen, P. M. (2020). Immunomodulatory Roles of Polysaccharide Capsules in the Intestine. *Frontiers in Immunology*, 11.
<https://doi.org/10.3389/fimmu.2020.00690>
- Ichiyama, T., Kuroda, K., Nagai, Y., Urushiyama, D., Ohno, M., Yamaguchi, T., Nagayoshi, M., Sakuraba, Y., Yamasaki, F., Hata, K., Miyamoto, S., Itakura, A., Takeda, S., & Tanaka, A. (2021). Analysis of vaginal and endometrial microbiota communities in infertile women with a history of repeated implantation failure. *Reproductive Medicine and Biology*, 20(3). <https://doi.org/10.1002/rmb2.12389>
- Iwami, N., Kawamata, M., Ozawa, N., Yamamoto, T., Watanabe, E., Mizuuchi, M., Moriwaka, O., & Kamiya, H. (2023). Therapeutic intervention based on gene sequencing analysis of microbial 16S ribosomal RNA of the intrauterine microbiome improves pregnancy outcomes in IVF patients: a prospective cohort study. *Journal of Assisted Reproduction and Genetics*, 40(1), 125–135.
<https://doi.org/10.1007/s10815-022-02688-6>

- Jarde, A., Lewis-Mikhael, A.-M., Moayyedi, P., Stearns, J. C., Collins, S. M., Beyene, J., & McDonald, S. D. (2018). Pregnancy outcomes in women taking probiotics or prebiotics: a systematic review and meta-analysis. *BMC Pregnancy and Childbirth*, 18(1), 14. <https://doi.org/10.1186/s12884-017-1629-5>
- Jarosik, G. P., Land, C. B., Duhon, P., Chandler, R., & Mercer, T. (1998). Acquisition of Iron by *Gardnerella vaginalis*. *Infection and Immunity*, 66(10), 5041–5047. <https://doi.org/10.1128/IAI.66.10.5041-5047.1998>
- Jensterle, M., Janez, A., Fliers, E., DeVries, J. H., Vrtacnik-Bokal, E., & Siegelaar, S. E. (2019). The role of glucagon-like peptide-1 in reproduction: from physiology to therapeutic perspective. *Human Reproduction Update*, 25(4), 504–517. <https://doi.org/10.1093/humupd/dmz019>
- Johnson, J. S., Spakowicz, D. J., Hong, B.-Y., Petersen, L. M., Demkowicz, P., Chen, L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., Sodergren, E., & Weinstock, G. M. (2019). Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications*, 10(1), 5029. <https://doi.org/10.1038/s41467-019-13036-1>
- Joseph, T. A., & Pe'er, I. (2021). *An Introduction to Whole-Metagenome Shotgun Sequencing Studies* (pp. 107–122). https://doi.org/10.1007/978-1-0716-1103-6_6
- Jung, H.-S., Ehlers, M. M., Lombaard, H., Redelinghuys, M. J., & Kock, M. M. (2017). Etiology of bacterial vaginosis and polymicrobial biofilm formation. *Critical Reviews in Microbiology*, 43(6), 651–667. <https://doi.org/10.1080/1040841X.2017.1291579>

- Juttukonda, L. J., & Skaar, E. P. (2015). Manganese homeostasis and utilization in pathogenic bacteria. *Molecular Microbiology*, 97(2), 216–228. <https://doi.org/10.1111/mmi.13034>
- Kadogami, D., Nakaoka, Y., & Morimoto, Y. (2020). Use of a vaginal probiotic suppository and antibiotics to influence the composition of the endometrial microbiota. *Reproductive Biology*, 20(3), 307–314. <https://doi.org/10.1016/j.repbio.2020.07.001>
- Khoa, L., Trieu Vy, N., Tai Loc, N., Thang, L., Anh, D., Hieu, N., Nhu, G., Tao, N., Hoang, L., & Handelsman, D. (2022). Antibiotic usage in surgical sperm retrievals among IVF centers. *Asian Journal of Andrology*, 24(6), 633. <https://doi.org/10.4103/aja2021123>
- Khodaparast, L., Khodaparast, L., Shahrooei, M., Stijlemans, B., Merckx, R., Baatsen, P., O'Gara, J. P., Waters, E., Van Mellaert, L., & Van Eldere, J. (2016). The Possible Role of Staphylococcus epidermidis LPxTG Surface Protein SesC in Biofilm Formation. *PLOS ONE*, 11(1), e0146704. <https://doi.org/10.1371/journal.pone.0146704>
- Khorchid, A., & Ikura, M. (2006). Bacterial histidine kinase as signal sensor and transducer. *The International Journal of Biochemistry & Cell Biology*, 38(3), 307–312. <https://doi.org/10.1016/j.biocel.2005.08.018>
- Kindinger, L. M., Bennett, P. R., Lee, Y. S., Marchesi, J. R., Smith, A., Cacciatore, S., Holmes, E., Nicholson, J. K., Teoh, T. G., & MacIntyre, D. A. (2017). The interaction between vaginal microbiota, cervical length, and vaginal progesterone treatment for preterm birth risk. *Microbiome*, 5(1), 6. <https://doi.org/10.1186/s40168-016-0223-9>

- Kitaya, K., Matsubayashi, H., Yamaguchi, K., Nishiyama, R., Takaya, Y., Ishikawa, T., Yasuo, T., & Yamada, H. (2016). Chronic Endometritis: Potential Cause of Infertility and Obstetric and Neonatal Complications. *American Journal of Reproductive Immunology*, 75(1), 13–22. <https://doi.org/10.1111/aji.12438>
- Kitaya, K., Nagai, Y., Arai, W., Sakuraba, Y., & Ishikawa, T. (2019). Characterization of Microbiota in Endometrial Fluid and Vaginal Secretions in Infertile Women with Repeated Implantation Failure. *Mediators of Inflammation*, 2019. <https://doi.org/10.1155/2019/4893437>
- Kitaya, K., & Yasuo, T. (2023). Commonalities and Disparities between Endometriosis and Chronic Endometritis: Therapeutic Potential of Novel Antibiotic Treatment Strategy against Ectopic Endometrium. *International Journal of Molecular Sciences*, 24(3), 2059. <https://doi.org/10.3390/ijms24032059>
- Králíčková, M., Fiala, L., Losan, P., Tomes, P., & Vetvicka, V. (2018). Altered Immunity in Endometriosis: What Came First? *Immunological Investigations*, 47(6), 569–582. <https://doi.org/10.1080/08820139.2018.1467926>
- Kunaseh, J., Waiyaput, W., Chanchaem, P., Sawaswong, V., Permpech, R., Payungporn, S., & Sophonsritsuk, A. (2022). Vaginal microbiome of women with adenomyosis: A case-control study. *PLOS ONE*, 17(2), e0263283. <https://doi.org/10.1371/journal.pone.0263283>
- Kurtz, Z. D., Müller, C. L., Miraldi, E. R., Littman, D. R., Blaser, M. J., & Bonneau, R. A. (2015). Sparse and Compositionally Robust Inference of Microbial Ecological Networks. *PLOS Computational Biology*, 11(5), e1004226. <https://doi.org/10.1371/journal.pcbi.1004226>

- Kyono, K., Hashimoto, T., Kikuchi, S., Nagai, Y., & Sakuraba, Y. (2018). A pilot study and case reports on endometrial microbiota and pregnancy outcome: An analysis using 16S rRNA gene sequencing among IVF patients, and trial therapeutic intervention for dysbiotic endometrium. *Reproductive Medicine and Biology*, 18(1), 72–82. <https://doi.org/10.1002/rmb2.12250>
- Kyono, K., Hashimoto, T., Nagai, Y., & Sakuraba, Y. (2018). Analysis of endometrial microbiota by 16S ribosomal RNA gene sequencing among infertile patients: a single-center pilot study. *Reproductive Medicine and Biology*, 17(3), 297–306. <https://doi.org/10.1002/rmb2.12105>
- Lagerström, M. C., & Schiöth, H. B. (2008). Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nature Reviews Drug Discovery*, 7(4), 339–357. <https://doi.org/10.1038/nrd2518>
- Lauder, A. P., Roche, A. M., Sherrill-Mix, S., Bailey, A., Laughlin, A. L., Bittinger, K., Leite, R., Elovitz, M. A., Parry, S., & Bushman, F. D. (2016). Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota. *Microbiome*, 4(1), 29. <https://doi.org/10.1186/s40168-016-0172-3>
- Laurence, M., Hatzis, C., & Brash, D. E. (2014). Common Contaminants in Next-Generation Sequencing That Hinder Discovery of Low-Abundance Microbes. *PLoS ONE*, 9(5), e97876. <https://doi.org/10.1371/journal.pone.0097876>
- Lederberg, J., & McCray, A. T. (2001). 'Ome Sweet 'Omics - A Genealogical Treasury of Words. The Scientist. <https://lhncbc.nlm.nih.gov/LHC-publications/pubs/OmeSweetOmicsAGenealogicalTreasuryofWords.html>

- Leoni, C., Ceci, O., Manzari, C., Fosso, B., Volpicella, M., Ferrari, A., Fiorella, P., Pesole, G., Cicinelli, E., & Ceci, L. R. (2019). Human Endometrial Microbiota at Term of Normal Pregnancies. *Genes*, 10(12), 971. <https://doi.org/10.3390/genes10120971>
- Lerminiaux, N. A., & Cameron, A. D. S. (2019). Horizontal transfer of antibiotic resistance genes in clinical environments. *Canadian Journal of Microbiology*, 65(1), 34–44. <https://doi.org/10.1139/cjm-2018-0275>
- Lessey, B. A., & Young, S. L. (2019). What exactly is endometrial receptivity? *Fertility and Sterility*, 111(4), 611–617. <https://doi.org/10.1016/j.fertnstert.2019.02.009>
- Lev-Sagie, A., Goldman-Wohl, D., Cohen, Y., Dori-Bachash, M., Leshem, A., Mor, U., Strahilevitz, J., Moses, A. E., Shapiro, H., Yagel, S., & Elinav, E. (2019). Vaginal microbiome transplantation in women with intractable bacterial vaginosis. *Nature Medicine*, 25(10), 1500–1504. <https://doi.org/10.1038/s41591-019-0600-6>
- Li, F., Chen, C., Wei, W., Wang, Z., Dai, J., Hao, L., Song, L., Zhang, X., Zeng, L., Du, H., Tang, H., Liu, N., Yang, H., Wang, J., Madsen, L., Brix, S., Kristiansen, K., Xu, X., Li, J., ... Jia, H. (2018). The metagenome of the female upper reproductive tract. *GigaScience*, 7(10). <https://doi.org/10.1093/gigascience/giy107>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, 26(5), 589–595. <https://doi.org/10.1093/bioinformatics/btp698>

- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Liang, J., Li, M., Zhang, L., Yang, Y., Jin, X., Zhang, Q., Lv, T., Huang, Z., Liao, Q., & Tong, X. (2023). Analysis of the microbiota composition in the genital tract of infertile patients with chronic endometritis or endometrial polyps. *Frontiers in Cellular and Infection Microbiology*, 13. <https://doi.org/10.3389/fcimb.2023.1125640>
- Lin, Q., Duan, H., Wang, S., Guo, Z., Wang, S., Chang, Y., Chen, C., Shen, M., Shou, H., & Zhou, C. (2023). Endometrial microbiota in women with and without adenomyosis: A pilot study. *Frontiers in Microbiology*, 14. <https://doi.org/10.3389/fmicb.2023.1075900>
- Lin, Z., Shi, J.-L., Chen, M., Zheng, Z.-M., Li, M.-Q., & Shao, J. (2023). CCL2: An important cytokine in normal and pathological pregnancies: A review. *Frontiers in Immunology*, 13. <https://doi.org/10.3389/fimmu.2022.1053457>
- Linjawi, S., Li, T. C., Tuckerman, E. M., Blakemore, A. I. F., & Laird, S. M. (2004). Expression of interleukin-11 receptor α and interleukin-11 protein in the endometrium of normal fertile women and women with recurrent miscarriage. *Journal of Reproductive Immunology*, 64(1–2), 145–155. <https://doi.org/10.1016/j.jri.2004.08.007>
- Liston, A., Carr, E. J., & Linterman, M. A. (2016). Shaping Variation in the Human Immune System. *Trends in Immunology*, 37(10), 637–646. <https://doi.org/10.1016/j.it.2016.08.002>

- Liu, F., Zhou, Y., Zhu, L., Wang, Z., Ma, L., He, Y., & Fu, P. (2021). Comparative metagenomic analysis of the vaginal microbiome in healthy women. *Synthetic and Systems Biotechnology*, 6(2), 77–84. <https://doi.org/10.1016/j.synbio.2021.04.002>
- Liu, F.-T., Yang, S., Yang, Z., Zhou, P., Peng, T., Yin, J., Ye, Z., Shan, H., Yu, Y., & Li, R. (2022). An Altered Microbiota in the Lower and Upper Female Reproductive Tract of Women with Recurrent Spontaneous Abortion. *Microbiology Spectrum*, 10(3). <https://doi.org/10.1128/spectrum.00462-22>
- Liu, X., Cao, Y., Xie, X., Qin, X., He, X., Shi, C., Zeng, W., Guo, Y., & Lin, Y. (2021). Association between vaginal microbiota and risk of early pregnancy miscarriage. *Comparative Immunology, Microbiology and Infectious Diseases*, 77, 101669. <https://doi.org/10.1016/j.cimid.2021.101669>
- Liu, Y., Wong, K. K.-W., Ko, E. Y.-L., Chen, X., Huang, J., Tsui, S. K.-W., Li, T. C., & Chim, S. S.-C. (2018). Systematic Comparison of Bacterial Colonization of Endometrial Tissue and Fluid Samples in Recurrent Miscarriage Patients: Implications for Future Endometrial Microbiome Studies. *Clinical Chemistry*, 64(12), 1743–1752. <https://doi.org/10.1373/clinchem.2018.289306>
- Locher, K. P. (2016). Mechanistic diversity in ATP-binding cassette (ABC) transporters. *Nature Structural & Molecular Biology*, 23(6), 487–493. <https://doi.org/10.1038/nsmb.3216>
- López-Moreno, A., & Aguilera, M. (2021). Vaginal Probiotics for Reproductive Health and Related Dysbiosis: Systematic Review and Meta-Analysis. *Journal of Clinical Medicine*, 10(7), 1461. <https://doi.org/10.3390/jcm10071461>

- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550. <https://doi.org/10.1186/s13059-014-0550-8>
- Lozano, F. M., Bernabeu, A., Lledo, B., Morales, R., Diaz, M., Aranda, F. I., Llacer, J., & Bernabeu, R. (2021). Characterization of the vaginal and endometrial microbiome in patients with chronic endometritis. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 263. <https://doi.org/10.1016/j.ejogrb.2021.05.045>
- Lozano, F. M., Lledó, B., Morales, R., Cascales, A., Hortal, M., Bernabeu, A., & Bernabeu, R. (2023). Characterization of the Endometrial Microbiome in Patients with Recurrent Implantation Failure. *Microorganisms*, 11(3), 741. <https://doi.org/10.3390/microorganisms11030741>
- Macklaim, J. M., Gloor, G. B., Anukam, K. C., Cribby, S., & Reid, G. (2011). At the crossroads of vaginal health and disease, the genome sequence of *Lactobacillus iners* AB-1. *Proceedings of the National Academy of Sciences*, 108(supplement_1), 4688–4695. <https://doi.org/10.1073/pnas.1000086107>
- Magoč, T., & Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, 27(21), 2957–2963. <https://doi.org/10.1093/bioinformatics/btr507>
- Manos, J. (2022). The human microbiome in disease and pathology. *APMIS*, 130(12), 690–705. <https://doi.org/10.1111/apm.13225>
- Marchesi, J. R., & Ravel, J. (2015). The vocabulary of microbiome research: a proposal. *Microbiome*, 3(1), 31. <https://doi.org/10.1186/s40168-015-0094-5>

- Matsuura, M. (2013). Structural Modifications of Bacterial Lipopolysaccharide that Facilitate Gram-Negative Bacteria Evasion of Host Innate Immunity. *Frontiers in Immunology*, 4. <https://doi.org/10.3389/fimmu.2013.00109>
- Mitchell, C. M., Haick, A., Nkwopara, E., Garcia, R., Rendi, M., Agnew, K., Fredricks, D. N., & Eschenbach, D. (2015). Colonization of the upper genital tract by vaginal bacterial species in nonpregnant women. *American Journal of Obstetrics and Gynecology*, 212(5), 611.e1-611.e9. <https://doi.org/10.1016/j.ajog.2014.11.043>
- Molina, N. M., Sola-Leyva, A., Haahr, T., Aghajanova, L., Laudanski, P., Castilla, J. A., & Altmäe, S. (2021). Analysing endometrial microbiome: methodological considerations and recommendations for good practice. *Human Reproduction*, 36(4), 859–879. <https://doi.org/10.1093/humrep/deab009>
- Mor, G., & Cardenas, I. (2010). The Immune System in Pregnancy: A Unique Complexity. *American Journal of Reproductive Immunology*, 63(6), 425–433. <https://doi.org/10.1111/j.1600-0897.2010.00836.x>
- Moreno, I., Capalbo, A., Mas, A., Garrido-Gomez, T., Roson, B., Poli, M., Dimitriadis, E., Santamaria, X., Vilella, F., & Simon, C. (2023). The human periconceptional maternal-embryonic space in health and disease. *Physiological Reviews*, 103(3), 1965–2038. <https://doi.org/10.1152/physrev.00050.2021>
- Moreno, I., Cicinelli, E., Garcia-Grau, I., Gonzalez-Monfort, M., Bau, D., Vilella, F., De Ziegler, D., Resta, L., Valbuena, D., & Simon, C. (2018). The diagnosis of chronic endometritis in infertile asymptomatic women: a comparative study of histology, microbial cultures, hysteroscopy, and molecular microbiology. *American Journal of Obstetrics and Gynecology*, 218(6), 602.e1-602.e16. <https://doi.org/10.1016/j.ajog.2018.02.012>

- Moreno, I., Codoñer, F. M., Vilella, F., Valbuena, D., Martinez-Blanch, J. F., Jimenez-Almazán, J., Alonso, R., Alamá, P., Remohí, J., Pellicer, A., Ramon, D., & Simon, C. (2016). Evidence that the endometrial microbiota has an effect on implantation success or failure. *American Journal of Obstetrics and Gynecology*, 215(6), 684–703. <https://doi.org/10.1016/j.ajog.2016.09.075>
- Moreno, I., Garcia-Grau, I., Bau, D., Perez-Villaroya, D., Gonzalez-Monfort, M., Vilella, F., Romero, R., & Simón, C. (2020). The first glimpse of the endometrial microbiota in early pregnancy. *American Journal of Obstetrics and Gynecology*, 222(4), 296–305. <https://doi.org/10.1016/j.ajog.2020.01.031>
- Moreno, I., Garcia-Grau, I., Perez-Villaroya, D., Gonzalez-, M., Bahçeci, M., Barrionuevo, M. J., Taguchi, S., Puente, E., Dimattina, M., Wei Lim, M., Meneghini, G., Aubuchon, M., Izquierdo, A., Perez-Olgati, M., Chavez, A., Bau, D., Gomez, C., Valbuena, D., Vilella, F., ... Fe, S. (2022). Endometrial microbiota composition is associated with reproductive outcome in infertile patients. *Microbiome*, 10(1). <https://doi.org/10.1101/2021.02.05.21251207>
- Moustafa, S., & Young, S. (2020). Diagnostic and therapeutic options in recurrent implantation failure. *F1000Research*, 9, 208. <https://doi.org/10.12688/f1000research.22403.1>
- Munita, J. M., & Arias, C. A. (2016). Mechanisms of Antibiotic Resistance. *Microbiology Spectrum*, 4(2). <https://doi.org/10.1128/microbiolspec.VMBF-0016-2015>
- Muraoka, A., Suzuki, M., Hamaguchi, T., Watanabe, S., Iijima, K., Murofushi, Y., Shinjo, K., Osuka, S., Hariyama, Y., Ito, M., Ohno, K., Kiyono, T., Kyo, S., Iwase, A., Kikkawa, F., Kajiyama, H., & Kondo, Y. (2023). *Fusobacterium* infection facilitates the development of endometriosis through the phenotypic transition of

- endometrial fibroblasts. *Science Translational Medicine*, 15(700).
<https://doi.org/10.1126/scitranslmed.add1531>
- Muzny, C. A., Łaniewski, P., Schwebke, J. R., & Herbst-Kralovetz, M. M. (2020). Host–vaginal microbiota interactions in the pathogenesis of bacterial vaginosis. *Current Opinion in Infectious Diseases*, 33(1), 59–65.
<https://doi.org/10.1097/QCO.0000000000000620>
- Nearing, J. T., Douglas, G. M., Hayes, M. G., MacDonald, J., Desai, D. K., Allward, N., Jones, C. M. A., Wright, R. J., Dhanani, A. S., Comeau, A. M., & Langille, M. G. I. (2022). Microbiome differential abundance methods produce different results across 38 datasets. *Nature Communications*, 13(1), 342.
<https://doi.org/10.1038/s41467-022-28034-z>
- Ng, S.-W., Norwitz, G. A., Pavlicev, M., Tilburgs, T., Simón, C., & Norwitz, E. R. (2020). Endometrial Decidualization: The Primary Driver of Pregnancy Health. *International Journal of Molecular Sciences*, 21(11), 4092.
<https://doi.org/10.3390/ijms21114092>
- Ni, N., & Li, Q. (2017). TGFβ superfamily signaling and uterine decidualization. *Reproductive Biology and Endocrinology*, 15(1), 84.
<https://doi.org/10.1186/s12958-017-0303-0>
- Nichols, R. G., & Davenport, E. R. (2021). The relationship between the gut microbiome and host gene expression: a review. *Human Genetics*, 140(5), 747–760. <https://doi.org/10.1007/s00439-020-02237-0>
- NIH HMP Working Group, Peterson, J., Garges, S., Giovanni, M., McInnes, P., Wang, L., Schloss, J. A., Bonazzi, V., McEwen, J. E., Wetterstrand, K. A., Deal, C., Baker, C. C., Di Francesco, V., Howcroft, T. K., Karp, R. W., Lunsford, R. D., Wellington,

- C. R., Belachew, T., Wright, M., ... Guyer, M. (2009). The NIH Human Microbiome Project. *Genome Research*, 19(12), 2317–2323.
<https://doi.org/10.1101/gr.096651.109>
- Novak, J., Ravel, J., Ma, B., Ferreira, C. S. T., Tristão, A. da R., Silva, M. G., & Marconi, C. (2022). Characteristics associated with *Lactobacillus iners* -dominated vaginal microbiota. *Sexually Transmitted Infections*, 98(5), 353–359.
<https://doi.org/10.1136/sextrans-2020-054824>
- Nutsch, K. M., & Hsieh, C.-S. (2012). T cell tolerance and immunity to commensal bacteria. *Current Opinion in Immunology*, 24(4), 385–391.
<https://doi.org/10.1016/j.coi.2012.04.009>
- Odendaal, J., Black, N., Bennett, P. R., Brosens, J., Quenby, S., & MacIntyre, D. A. (2024). The endometrial microbiota and early pregnancy loss. *Human Reproduction*. <https://doi.org/10.1093/humrep/dead274>
- Ojala, T., Kankuri, E., & Kankainen, M. (2023). Understanding human health through metatranscriptomics. *Trends in Molecular Medicine*, 29(5), 376–389.
<https://doi.org/10.1016/j.molmed.2023.02.002>
- Orelle, C., Mathieu, K., & Jault, J.-M. (2019). Multidrug ABC transporters in bacteria. *Research in Microbiology*, 170(8), 381–391.
<https://doi.org/10.1016/j.resmic.2019.06.001>
- Pantos, K., Grigoriadis, S., Maziotis, E., Pistola, K., Xystra, P., Pantou, A., Kokkali, G., Pappas, A., Lambropoulou, M., Sfakianoudis, K., & Simopoulou, M. (2022). The Role of Interleukins in Recurrent Implantation Failure: A Comprehensive Review of the Literature. *International Journal of Molecular Sciences*, 23(4), 2198.
<https://doi.org/10.3390/ijms23042198>

- Park, C. H., Hong, C., Lee, A., Sung, J., & Hwang, T. H. (2022). Multi-omics reveals microbiome, host gene expression, and immune landscape in gastric carcinogenesis. *IScience*, 25(3), 103956. <https://doi.org/10.1016/j.isci.2022.103956>
- Parvanov, D., Ganeva, R., Ruseva, M., Handzhiyska, M., Tihomirova, T., Chapanova, S., Staneva, R., Rukova, B., Pancheva, M., Serafimova, M., Metodiev, D., Stamenov, G., & Hadjidekova, S. (2023). Association between endometrial microbiome and implantation success in women with frozen embryo transfer: results of a prospective cohort study. *Biotechnology & Biotechnological Equipment*, 37(1). <https://doi.org/10.1080/13102818.2023.2250007>
- Paula, R., Oliani, A. H., Vaz-Oliani, D. C. M., D'Ávila, S. C. G. P., Oliani, S. M., & Gil, C. D. (2015). The intricate role of mast cell proteases and the annexin A1-FPR1 system in abdominal wall endometriosis. *Journal of Molecular Histology*, 46(1), 33–43. <https://doi.org/10.1007/s10735-014-9595-y>
- Pelzer, E. S., Willner, D., Buttini, M., & Huygens, F. (2018). A role for the endometrial microbiome in dysfunctional menstrual bleeding. *Antonie van Leeuwenhoek*, 111(6), 933–943. <https://doi.org/10.1007/s10482-017-0992-6>
- Pereira, N., Hutchinson, A. P., Lekovich, J. P., Hobeika, E., & Elias, R. T. (2016). Antibiotic Prophylaxis for Gynecologic Procedures prior to and during the Utilization of Assisted Reproductive Technologies: A Systematic Review. *Journal of Pathogens*, 2016. <https://doi.org/10.1155/2016/4698314>
- Perez-Muñoz, M. E., Arrieta, M.-C., Ramer-Tait, A. E., & Walter, J. (2017). A critical assessment of the “sterile womb” and “in utero colonization” hypotheses:

- implications for research on the pioneer infant microbiome. *Microbiome*, 5(1), 48.
<https://doi.org/10.1186/s40168-017-0268-4>
- Petrova, M. I., Lievens, E., Malik, S., Imholz, N., & Lebeer, S. (2015). Lactobacillus species as biomarkers and agents that can promote various aspects of vaginal health. *Frontiers in Physiology*, 6. <https://doi.org/10.3389/fphys.2015.00081>
- Pi, H., & Helmann, J. D. (2017). Ferrous iron efflux systems in bacteria. *Metallomics*, 9(7), 840–851. <https://doi.org/10.1039/C7MT00112F>
- Piccinni, M.-P., Raghupathy, R., Saito, S., & Szekeres-Bartho, J. (2021). Cytokines, Hormones and Cellular Regulatory Mechanisms Favoring Successful Reproduction. *Frontiers in Immunology*, 12. <https://doi.org/10.3389/fimmu.2021.717808>
- Pirih, N., & Kunej, T. (2018). An Updated Taxonomy and a Graphical Summary Tool for Optimal Classification and Comprehension of Omics Research. *OMICS: A Journal of Integrative Biology*, 22(5), 337–353. <https://doi.org/10.1089/omi.2017.0186>
- Potempa, J., & Pike, R. N. (2009). Corruption of Innate Immunity by Bacterial Proteases. *Journal of Innate Immunity*, 1(2), 70–87. <https://doi.org/10.1159/000181144>
- Prasad, Uv., Vasu, D., Gowtham, Rr., Pradeep, K., Swarupa, V., Yeswanth, S., Choudhary, A., & Sarma, P. V. G. K. (2017). Cloning, Expression and Characterization of NAD Kinase from Staphylococcus aureus Involved in the Formation of NADP (H): A Key Molecule in the Maintaining of Redox Status and Biofilm Formation. *Advanced Biomedical Research*, 6(1), 97. <https://doi.org/10.4103/2277-9175.211833>

- Pruitt, K. D. (2005). NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research*, 33(Database issue), D501–D504. <https://doi.org/10.1093/nar/gki025>
- Punzón-Jiménez, P., & Labarta, E. (2021). The impact of the female genital tract microbiome in women health and reproduction: a review. *Journal of Assisted Reproduction and Genetics*. <https://doi.org/10.1007/s10815-021-02247-5>
- Qi, R., Zhang, T., Zhang, Y., Chung, J. P. W., Yang, W.-J., & Wang, C. C. (2023). Association of angiotensin II and receptors in peri-implantation endometrium with microvessel density and pregnancy outcomes of women with recurrent implantation failure after embryo transfer. *Frontiers in Endocrinology*, 14. <https://doi.org/10.3389/fendo.2023.1206326>
- Quenby, S., Nik, H., Innes, B., Lash, G., Turner, M., Drury, J., & Bulmer, J. (2008). Uterine natural killer cells and angiogenesis in recurrent reproductive failure. *Human Reproduction*, 24(1), 45–54. <https://doi.org/10.1093/humrep/den348>
- R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S., & Medzhitov, R. (2004). Recognition of Commensal Microflora by Toll-Like Receptors Is Required for Intestinal Homeostasis. *Cell*, 118(2), 229–241. <https://doi.org/10.1016/j.cell.2004.07.002>
- Rather, M. A., Gupta, K., & Mandal, M. (2021). Microbial biofilm: formation, architecture, antibiotic resistance, and control strategies. *Brazilian Journal of Microbiology*, 52(4), 1701–1718. <https://doi.org/10.1007/s42770-021-00624-x>

- Ravel, J., Gajer, P., Abdo, Z., Schneider, G. M., Koenig, S. S. K., McCulle, S. L., Karlebach, S., Gorle, R., Russell, J., Tacket, C. O., Brotman, R. M., Davis, C. C., Ault, K., Peralta, L., & Forney, L. J. (2011). Vaginal microbiome of reproductive-age women. *Proceedings of the National Academy of Sciences*, 108(Suppl 1), 4680–4687. <https://doi.org/10.1073/pnas.1002611107>
- Ravel, J., Moreno, I., & Simón, C. (2021). Bacterial vaginosis and its association with infertility, endometritis, and pelvic inflammatory disease. *American Journal of Obstetrics and Gynecology*, 224(3), 251–257. <https://doi.org/10.1016/j.ajog.2020.10.019>
- Recine, N., Palma, E., Domenici, L., Giorgini, M., Imperiale, L., Sassu, C., Musella, A., Marchetti, C., Muzii, L., & Benedetti Panici, P. (2016). Restoring vaginal microbiota: biological control of bacterial vaginosis. A prospective case–control study using *Lactobacillus rhamnosus* BMX 54 as adjuvant treatment against bacterial vaginosis. *Archives of Gynecology and Obstetrics*, 293(1), 101–107. <https://doi.org/10.1007/s00404-015-3810-2>
- Riganelli, L., Iebba, V., Piccioni, M., Illuminati, I., Bonfiglio, G., Neroni, B., Calvo, L., Gagliardi, A., Levrero, M., Merlino, L., Mariani, M., Capri, O., Pietrangeli, D., Schippa, S., & Guerrieri, F. (2020). Structural Variations of Vaginal and Endometrial Microbiota: Hints on Female Infertility. *Frontiers in Cellular and Infection Microbiology*, 10(350). <https://doi.org/10.3389/fcimb.2020.00350>
- Rinninella, E., Raoul, P., Cintoni, M., Franceschi, F., Miggiano, G., Gasbarrini, A., & Mele, M. (2019). What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. *Microorganisms*, 7(1), 14. <https://doi.org/10.3390/microorganisms7010014>

- Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1), 139–140. <https://doi.org/10.1093/bioinformatics/btp616>
- Sabbatini, S., Visconti, S., Gentili, M., Lusenti, E., Nunzi, E., Ronchetti, S., Perito, S., Gaziano, R., & Monari, C. (2021). Lactobacillus iners Cell-Free Supernatant Enhances Biofilm Formation and Hyphal/Pseudohyphal Growth by Candida albicans Vaginal Isolates. *Microorganisms*, 9(12), 2577. <https://doi.org/10.3390/microorganisms9122577>
- Salas-Jara, M., Ilabaca, A., Vega, M., & García, A. (2016). Biofilm Forming Lactobacillus: New Challenges for the Development of Probiotics. *Microorganisms*, 4(3), 35. <https://doi.org/10.3390/microorganisms4030035>
- Sallam, H. N., Gelbaya, T. A., Maldonado Rosas, I., Anagnostopoulou, C., Sallam, N., & Agarwal, A. (2022). Clinical aspects of oocyte retrieval and embryo transfer: tips and tricks for the novice and the expert. *Panminerva Medica*, 64(2). <https://doi.org/10.23736/S0031-0808.22.04679-1>
- Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., Turner, P., Parkhill, J., Loman, N. J., & Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology*, 12, 87. <https://doi.org/10.1186/S12915-014-0087-Z>
- Sauer, K., Stoodley, P., Goeres, D. M., Hall-Stoodley, L., Burmølle, M., Stewart, P. S., & Bjarnsholt, T. (2022). The biofilm life cycle: expanding the conceptual model of biofilm formation. *Nature Reviews Microbiology*, 20(10), 608–620. <https://doi.org/10.1038/s41579-022-00767-0>

- Savicheva, A. M. (2023). Molecular Testing for the Diagnosis of Bacterial Vaginosis. *International Journal of Molecular Sciences*, 25(1), 449. <https://doi.org/10.3390/ijms25010449>
- Schindler, Y., Rahav, G., Nissan, I., Valenci, G., Ravins, M., Hanski, E., Ment, D., Tekes-Manova, D., & Maor, Y. (2023). Type VII secretion system and its effect on group B Streptococcus virulence in isolates obtained from newborns with early onset disease and colonized pregnant women. *Frontiers in Cellular and Infection Microbiology*, 13. <https://doi.org/10.3389/fcimb.2023.1168530>
- Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., & Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. *Genome Biology*, 12(6), R60. <https://doi.org/10.1186/gb-2011-12-6-r60>
- Semmes, E. C., & Coyne, C. B. (2022). Innate immune defenses at the maternal-fetal interface. *Current Opinion in Immunology*, 74, 60–67. <https://doi.org/10.1016/j.coi.2021.10.007>
- Shannon, B., Gajer, P., Yi, T. J., Ma, B., Humphrys, M. S., Thomas-Pavanel, J., Chieza, L., Janakiram, P., Saunders, M., Tharao, W., Huibner, S., Shahabi, K., Ravel, J., & Kaul, R. (2017). Distinct Effects of the Cervicovaginal Microbiota and Herpes Simplex Type 2 Infection on Female Genital Tract Immunology. *The Journal of Infectious Diseases*, 215(9), 1366–1375. <https://doi.org/10.1093/infdis/jix088>
- Sharkey, L. K. R., & O'Neill, A. J. (2018). Antibiotic Resistance ABC-F Proteins: Bringing Target Protection into the Limelight. *ACS Infectious Diseases*, 4(3), 239–246. <https://doi.org/10.1021/acsinfecdis.7b00251>

- Sharma, D., Misba, L., & Khan, A. U. (2019). Antibiotics versus biofilm: an emerging battleground in microbial communities. *Antimicrobial Resistance & Infection Control*, 8(1), 76. <https://doi.org/10.1186/s13756-019-0533-3>
- Singh, S., Verma, N., & Taneja, N. (2019). The human gut resistome: Current concepts & future prospects. *Indian Journal of Medical Research*, 150(4), 345. https://doi.org/10.4103/ijmr.IJMR_1979_17
- Sirota, I., Zarek, S., & Segars, J. (2014). Potential Influence of the Microbiome on Infertility and Assisted Reproductive Technology. *Seminars in Reproductive Medicine*, 32(1), 35–42. <https://doi.org/10.1055/s-0033-1361821>
- Sola-Leyva, A., Andrés-León, E., Molina, N. M., Terron-Camero, L. C., Plaza-Díaz, J., Sáez-Lara, M. J., Gonzalvo, M. C., Sánchez, R., Ruíz, S., Martínez, L., & Altmäe, S. (2021). Mapping the entire functionally active endometrial microbiota. *Human Reproduction*, 36(4). <https://doi.org/10.1093/humrep/deaa372>
- Souza, S. V. de, Monteiro, P. B., Moura, G. A. de, Santos, N. O., Fontanezi, C. T. B., Gomes, I. de A., & Teixeira, C. A. (2023). Vaginal microbioma and the presence of *Lactobacillus* spp. as interferences in female fertility: A review system. *JBRA Assisted Reproduction*, 27. <https://doi.org/10.5935/1518-0557.20230006>
- Sparks, R. A., Purrier, B. G., Watt, P. J., & Elstein, M. (1981). Bacteriological colonisation of uterine cavity: role of tailed intrauterine contraceptive device. *BMJ*, 282(6271), 1189–1191. <https://doi.org/10.1136/bmj.282.6271.1189>
- Sprague, R., Kim, J. W., Kirimlioglu, E., Guo, X., Günay, N., Guzeloglu-Kayisli, O., Ozmen, A., Schatz, F., Imudia, A. N., Lockwood, C. J., Magness, R. R., & Kayisli, U. A. (2021). Catecholestradiol Activation of Adrenergic Receptors Induces Endometrial Cell Survival via p38 MAPK Signaling. *The Journal of Clinical*

- Endocrinology & Metabolism*, 106(2), 337–350.
<https://doi.org/10.1210/clinem/dgaa866>
- Stinson, L. F., Keelan, J. A., & Payne, M. S. (2019). Identification and removal of contaminating microbial <scp>DNA</scp> from <scp>PCR</scp> reagents: impact on low-biomass microbiome analyses. *Letters in Applied Microbiology*, 68(1), 2–8. <https://doi.org/10.1111/lam.13091>
- Su, X., Gao, Y., & Yang, R. (2022). Gut Microbiota-Derived Tryptophan Metabolites Maintain Gut and Systemic Homeostasis. *Cells*, 11(15), 2296. <https://doi.org/10.3390/cells11152296>
- Svenstrup, H. F., Fedder, J., Abraham-Peskir, J., Birkelund, S., & Christiansen, G. (2003). Mycoplasma genitalium attaches to human spermatozoa. *Human Reproduction (Oxford, England)*, 18(10), 2103–2109. <https://doi.org/10.1093/humrep/deg392>
- Swidsinski, A., Verstraelen, H., Loening-Baucke, V., Swidsinski, S., Mendling, W., & Halwani, Z. (2013). Presence of a Polymicrobial Endometrial Biofilm in Patients with Bacterial Vaginosis. *PLoS ONE*, 8(1), e53997. <https://doi.org/10.1371/journal.pone.0053997>
- Takamoto, N., Zhao, B., Tsai, S. Y., & DeMayo, F. J. (2002). Identification of Indian Hedgehog as a Progesterone-Responsive Gene in the Murine Uterus. *Molecular Endocrinology*, 16(10), 2338–2348. <https://doi.org/10.1210/me.2001-0154>
- Takimoto, K., Yamada, H., Shimada, S., Fukushi, Y., & Wada, S. (2023). Chronic Endometritis and Uterine Endometrium Microbiota in Recurrent Implantation Failure and Recurrent Pregnancy Loss. *Biomedicines*, 11(9), 2391. <https://doi.org/10.3390/biomedicines11092391>

- Tanner, M. A., Goebel, B. M., Dojka, M. A., & Pace, N. R. (1998). Specific Ribosomal DNA Sequences from Diverse Environmental Settings Correlate with Experimental Contaminants. *Applied and Environmental Microbiology*, 64(8), 3110–3113. <https://doi.org/10.1128/AEM.64.8.3110-3113.1998>
- Tapmeier, T. T., Rahmioglu, N., Lin, J., De Leo, B., Obendorf, M., Raveendran, M., Fischer, O. M., Bafligil, C., Guo, M., Harris, R. A., Hess-Stumpp, H., Laux-Biehlmann, A., Lowy, E., Lunter, G., Malzahn, J., Martin, N. G., Martinez, F. O., Manek, S., Mesch, S., ... Zondervan, K. T. (2021). Neuropeptide S receptor 1 is a nonhormonal treatment target in endometriosis. *Science Translational Medicine*, 13(608). <https://doi.org/10.1126/scitranslmed.abd6469>
- The Integrative HMP (iHMP) Research Network Consortium. (2014). The Integrative Human Microbiome Project: Dynamic Analysis of Microbiome-Host Omics Profiles during Periods of Human Health and Disease. *Cell Host & Microbe*, 16(3), 276–289. <https://doi.org/10.1016/j.chom.2014.08.014>
- Tomás, M., Palmeira-de-Oliveira, A., Simões, S., Martinez-de-Oliveira, J., & Palmeira-de-Oliveira, R. (2020). Bacterial vaginosis: Standard treatments and alternative strategies. *International Journal of Pharmaceutics*, 587, 119659. <https://doi.org/10.1016/j.ijpharm.2020.119659>
- Tortelli, B. A., Lewis, A. L., & Fay, J. C. (2021). The structure and diversity of strain-level variation in vaginal bacteria. *Microbial Genomics*, 7(3). <https://doi.org/10.1099/mgen.0.000543>
- Toson, B., Simon, C., & Moreno, I. (2022). The Endometrial Microbiome and Its Impact on Human Conception. *International Journal of Molecular Sciences*, 23(1), 485. <https://doi.org/10.3390/ijms23010485>

- Totorikaguena, L., Olabarrieta, E., Matorras, R., Alonso, E., Agirregoitia, E., & Agirregoitia, N. (2017). Mu opioid receptor in the human endometrium: dynamics of its expression and localization during the menstrual cycle. *Fertility and Sterility*, 107(4), 1070-1077.e1. <https://doi.org/10.1016/j.fertnstert.2017.01.020>
- Tuniyazi, M., & Zhang, N. (2023). Possible Therapeutic Mechanisms and Future Perspectives of Vaginal Microbiota Transplantation. *Microorganisms*, 11(6), 1427. <https://doi.org/10.3390/microorganisms11061427>
- Väinämö, S., Saqib, S., Kalliala, I., Kervinen, K., Lairo, K., Niinimäki, M., Halttunen-Nieminen, M., Virtanen, S., Nieminen, P., Salonen, A., & Holster, T. (2023). Longitudinal analysis of vaginal microbiota during IVF fresh embryo transfer and in early pregnancy. *Microbiology Spectrum*, 11(6). <https://doi.org/10.1128/spectrum.01650-23>
- van Houdt, R., Ma, B., Bruisten, S. M., Speksnijder, A. G. C. L., Ravel, J., & de Vries, H. J. C. (2018). *Lactobacillus iners* -dominated vaginal microbiota is associated with increased susceptibility to *Chlamydia trachomatis* infection in Dutch women: a case–control study. *Sexually Transmitted Infections*, 94(2), 117–123. <https://doi.org/10.1136/sextrans-2017-053133>
- Vitagliano, A., Saccardi, C., Noventa, M., Di Spiezio Sardo, A., Saccone, G., Cicinelli, E., Pizzi, S., Andrisani, A., & Litta, P. S. (2018). Effects of chronic endometritis therapy on in vitro fertilization outcome in women with repeated implantation failure: a systematic review and meta-analysis. *Fertility and Sterility*, 110(1), 103-112.e1. <https://doi.org/10.1016/j.fertnstert.2018.03.017>
- Wagner, S., Grin, I., Malmshemer, S., Singh, N., Torres-Vargas, C. E., & Westerhausen, S. (2018). Bacterial type III secretion systems: a complex device

- for the delivery of bacterial effector proteins into eukaryotic host cells. *FEMS Microbiology Letters*, 365(19). <https://doi.org/10.1093/femsle/fny201>
- Wang, J., Li, Z., Ma, X., Du, L., Jia, Z., Cui, X., Yu, L., Yang, J., Xiao, L., Zhang, B., Fan, H., & Zhao, F. (2021). Translocation of vaginal microbiota is involved in impairment and protection of uterine health. *Nature Communications*, 12(1), 4191. <https://doi.org/10.1038/s41467-021-24516-8>
- Wang, J., Zhu, N., Su, X., Gao, Y., & Yang, R. (2023). Gut-Microbiota-Derived Metabolites Maintain Gut and Systemic Immune Homeostasis. *Cells*, 12(5), 793. <https://doi.org/10.3390/cells12050793>
- Wang, W., Vilella, F., Alama, P., Moreno, I., Mignardi, M., Isakova, A., Pan, W., Simon, C., & Quake, S. R. (2020). Single-cell transcriptomic atlas of the human endometrium during the menstrual cycle. *Nature Medicine*, 26(10), 1644–1653. <https://doi.org/10.1038/s41591-020-1040-z>
- Wang, Z., & Zhao, Y. (2018). Gut microbiota derived metabolites in cardiovascular health and disease. *Protein & Cell*, 9(5), 416–431. <https://doi.org/10.1007/s13238-018-0549-0>
- Wee, B. A., Thomas, M., Sweeney, E. L., Frentiu, F. D., Samios, M., Ravel, J., Gajer, P., Myers, G., Timms, P., Allan, J. A., & Huston, W. M. (2018). A retrospective pilot study to determine whether the reproductive tract microbiota differs between women with a history of infertility and fertile women. *Australian and New Zealand Journal of Obstetrics and Gynaecology*, 58(3), 341–348. <https://doi.org/10.1111/ajo.12754>
- Wei, T., Simko, V. R., Levy, M., Xie, Y., Jin, Y., & Zemla, J. (2021). package “corrplot”: Visualization of a Correlation Matrix. 2017. *Version 0.84*.

- Wei, W., Zhang, X., Tang, H., Zeng, L., & Wu, R. (2020). Microbiota composition and distribution along the female reproductive tract of women with endometriosis. *Annals of Clinical Microbiology and Antimicrobials*, 19(1), 15. <https://doi.org/10.1186/s12941-020-00356-0>
- Wensel, C. R., Pluznick, J. L., Salzberg, S. L., & Sears, C. L. (2022). Next-generation sequencing: insights to advance clinical investigations of the microbiome. *Journal of Clinical Investigation*, 132(7). <https://doi.org/10.1172/JCI154944>
- Weyrich, L. S., Farrer, A. G., Eisenhofer, R., Arriola, L. A., Young, J., Selway, C. A., Handsley-Davis, M., Adler, C. J., Breen, J., & Cooper, A. (2019). Laboratory contamination over time during low-biomass sample analysis. *Molecular Ecology Resources*, 19(4), 982–996. <https://doi.org/10.1111/1755-0998.13011>
- Wolff, M. v. (2000). Regulated expression of cytokines in human endometrium throughout the menstrual cycle: dysregulation in habitual abortion. *Molecular Human Reproduction*, 6(7), 627–634. <https://doi.org/10.1093/molehr/6.7.627>
- Wu, F., Chen, X., Liu, Y., Liang, B., Xu, H., Li, T. C., & Wang, C. C. (2018). Decreased MUC1 in endometrium is an independent receptivity marker in recurrent implantation failure during implantation window. *Reproductive Biology and Endocrinology*, 16(1), 60. <https://doi.org/10.1186/s12958-018-0379-1>
- Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X., & Yu, G. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, 2(3), 100141. <https://doi.org/10.1016/j.xinn.2021.100141>

- Xu, X., Cai, X., Liu, X., & Guo, S.-W. (2021). Possible involvement of neuropeptide and neurotransmitter receptors in Adenomyosis. *Reproductive Biology and Endocrinology*, 19, 25. <https://doi.org/10.1186/s12958-021-00711-6>
- Yang, X., Tian, Y., Zheng, L., Luu, T., & Kwak-Kim, J. (2022). The Update Immune-Regulatory Role of Pro- and Anti-Inflammatory Cytokines in Recurrent Pregnancy Losses. *International Journal of Molecular Sciences*, 24(1), 132. <https://doi.org/10.3390/ijms24010132>
- Yeoman, C. J., Yildirim, S., Thomas, S. M., Durkin, A. S., Torralba, M., Sutton, G., Buhay, C. J., Ding, Y., Dugan-Rocha, S. P., Muzny, D. M., Qin, X., Gibbs, R. A., Leigh, S. R., Stumpf, R., White, B. A., Highlander, S. K., Nelson, K. E., & Wilson, B. A. (2010). Comparative Genomics of *Gardnerella vaginalis* Strains Reveals Substantial Differences in Metabolic and Virulence Potential. *PLoS ONE*, 5(8), e12411. <https://doi.org/10.1371/journal.pone.0012411>
- Yin, A., Guan, X., Zhang, J. V., & Niu, J. (2022a). Focusing on the role of secretin/adhesion (Class B) G protein-coupled receptors in placental development and preeclampsia. *Frontiers in Cell and Developmental Biology*, 10. <https://doi.org/10.3389/fcell.2022.959239>
- Yin, A., Guan, X., Zhang, J. V., & Niu, J. (2022b). Focusing on the role of secretin/adhesion (Class B) G protein-coupled receptors in placental development and preeclampsia. *Frontiers in Cell and Developmental Biology*, 10. <https://doi.org/10.3389/fcell.2022.959239>
- Zegers-Hochschild, F., Adamson, G. D., Dyer, S., Racowsky, C., de Mouzon, J., Sokol, R., Rienzi, L., Sunde, A., Schmidt, L., Cooke, I. D., Simpson, J. L., & van der Poel,

- S. (2017). The International Glossary on Infertility and Fertility Care, 2017. *Human Reproduction*, 32(9), 1786–1801. <https://doi.org/10.1093/humrep/dex234>
- Zeng, H., He, D., Hu, L., Abdullah, R. K., Zhang, L., Jiang, B., Xie, H., & Liu, N. (2022). Non-Lactobacillus dominance of the vagina is associated with reduced live birth rate following IVF/ICSI: a propensity score-matched cohort study. *Archives of Gynecology and Obstetrics*, 305(2), 519–528. <https://doi.org/10.1007/s00404-021-06171-y>
- Zhai, J., Ma, L., Chang, Z., & Yu, T. (2022). Increased expression of prokineticin 2 and its receptor in endometrium of recurrent implantation failure patients decreased the expression of MMP9 important for decidualization. *Reproductive Biology and Endocrinology*, 20(1), 76. <https://doi.org/10.1186/s12958-022-00947-w>
- Zhai, W., Wu, F., Zhang, Y., Fu, Y., & Liu, Z. (2019). The Immune Escape Mechanisms of Mycobacterium Tuberculosis. *International Journal of Molecular Sciences*, 20(2), 340. <https://doi.org/10.3390/ijms20020340>
- Zhang, Y., Thompson, K. N., Branck, T., Yan Yan, Nguyen, L. H., Franzosa, E. A., & Huttenhower, C. (2021). Metatranscriptomics for the Human Microbiome and Microbial Community Functional Profiling. *Annual Review of Biomedical Data Science*, 4(1), 279–311. <https://doi.org/10.1146/annurev-biodatasci-031121-103035>
- Zhao, L., Liu, Q., Huang, Q., Liu, F., Liu, H., & Wang, G. (2021). Isocitrate dehydrogenase of *Bacillus cereus* is involved in biofilm formation. *World Journal of Microbiology and Biotechnology*, 37(12), 207. <https://doi.org/10.1007/s11274-021-03175-3>

- Zheng, D., Liwinski, T., & Elinav, E. (2020). Interaction between microbiota and immunity in health and disease. *Cell Research*, 30(6), 492–506. <https://doi.org/10.1038/s41422-020-0332-7>
- Zheng, N., Guo, R., Wang, J., Zhou, W., & Ling, Z. (2021). Contribution of Lactobacillus iners to Vaginal Health and Diseases: A Systematic Review. *Frontiers in Cellular and Infection Microbiology*, 11. <https://doi.org/10.3389/fcimb.2021.792787>
- Zhou, X., Brown, C. J., Abdo, Z., Davis, C. C., Hansmann, M. A., Joyce, P., Foster, J. A., & Forney, L. J. (2007). Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women. *The ISME Journal*, 1(2), 121–133. <https://doi.org/10.1038/ismej.2007.12>
- Złotkowska, A., & Andronowska, A. (2019). Chemokines as the modulators of endometrial epithelial cells remodelling. *Scientific Reports*, 9(1), 12968. <https://doi.org/10.1038/s41598-019-49502-5>
- Zou, Y., Liu, X., Chen, P., Wang, Y., Li, W., & Huang, R. (2023). The endometrial microbiota profile influenced pregnancy outcomes in patients with repeated implantation failure: a retrospective study. *Journal of Reproductive Immunology*, 155, 103782. <https://doi.org/10.1016/j.jri.2022.103782>

9. ANNEXES

Annex 1. Virulence factors differentially abundant in NP compared to LB.

Category	VF gene name	Description	BaseMean	log2FC	p-adj
Effector delivery system (VFC0086)	<i>lirB</i>	Dot/Icm type IV secretion system effector LirB Dot/Icm T4SS secreted effectors (VF0798)	11.279	2.033	0.035
Exotoxin (VFC0235)	<i>ply</i>	pneumolysin Pneumolysin (VF0149)	11.781	2.264	0.035
	<i>HMPREF0424_RS00505</i>	thiol-activated cytolysin family protein Vaginolysin (VLY) (VF1316)	6.586	2.33	0.032
	<i>sly</i>	hemolysin Suilysin (VF1062)	5.755	2.436	0.02
		cholesterol-dependent cytolysin suilysin Pneumolysin (VF0149)	5.51	2.524	0.019
Immune modulation (VFC0258)	<i>lgtA</i>	N-acetylglucosamine glycosyltransferase LOS (VF0044)	2.915	2.006	0.035
	<i>wlaN</i>	beta-1;3 galactosyltransferase LOS (VF0326)	4.136	2.082	0.035
	<i>BCE_RS25805</i>	LytR family transcriptional regulator Polysaccharide capsule (VF0659)	13.753	2.092	0.05
	<i>wecB</i>	UDP-N-acetylglucosamine 2-epimerase (non-hydrolyzing) LPS (VF0085)	12.966	2.195	0.041
	<i>wbjD/wecB</i>	UDP-N-acetylglucosamine 2-epimerase (non-hydrolyzing)	13.444	2.248	0.035
		Capsular polysaccharide (VF0624)	13.444	2.248	0.035
Regulation (VFC0301)	<i>senX3</i>	two-component sensor histidine kinase SenX3 (VF0857)	13.601	2.099	0.049
			13.847	2.117	0.047
			13.753	2.141	0.043
			12.898	2.204	0.035
		two-component system sensor histidine kinase SenX3 SenX3 (VF0857)	13.409	2.117	0.046
		sensor histidine kinase SenX3 (VF0857)	13.488	2.117	0.044
		two component sensor histidine kinase SenX3 SenX3 (VF0857)	13.354	2.123	0.046
		histidine kinase SenX3 (VF0857)	12.578	2.178	0.035
Stress survival (VFC0282)	<i>bsh</i>	bile salt hydrolase BSH (VF0350)	5.822	2.121	0.035
			5.822	2.121	0.035
			5.808	2.121	0.035
			5.808	2.121	0.035
			5.808	2.121	0.035
			5.808	2.121	0.035
			5.808	2.121	0.035
			5.937	2.152	0.035
			9.463	2.163	0.043
			6.266	2.212	0.035

Annex 2. Unique virulence factors in the no pregnancy (NP) group.

Category	VF gene name	Description
Adherence (VFC0001)	<i>aatA</i>	autotransporter outer membrane beta-barrel domain-containing protein AatA, AIDA-I type (VF1133)
	<i>ABAYE_RS02125</i>	prepilin-type N-terminal cleavage/methylation domain-containing protein TFP (VF1334)
	<i>ABD1_RS15725</i>	prepilin-type N-terminal cleavage/methylation domain-containing protein TFP (VF1334)
	<i>ACI49667</i>	putative pilus tip protein PilB-type pili (PGS3) (VF0745)
	<i>acm</i>	collagen adhesin precursor Acn Acn (VF0419)
	<i>AvisC_010100012015</i>	surface-anchored fimbrial subunit Type 2 pili (VF1201)
	<i>AvisC_010100012020</i>	surface-anchored fimbrial subunit Type 2 pili (VF1201)
	<i>BC_RS12555</i>	choice-of-anchor A family protein Bcp pili (VF0669)
	<i>bcfD</i>	fimbrial adhesin protein Bcf (VF0950)
	<i>bcfG</i>	fimbrial chaparone Bcf (VF0950)
	<i>boaB</i>	autotransporter protein BoaB (VF0435)
	<i>cbpA</i>	collagen-binding protein A precursor CbpA (VF1203)
	<i>chpE</i>	probable chemotaxis protein Type IV pili (VF0082)
	<i>cna</i>	collagen adhesin precursor CNA (VF0005)
	<i>cotC</i>	CotC Adhesive fimbriae (VF0213)
	<i>cpa</i>	fibronectin-binding protein, GAS pilus tip adhesin Collagen binding protein (VF1039)
	<i>cpbA</i>	hypothetical protein Collagen binding protein (VF1039)
	<i>cwp66</i>	cell wall-binding protein Cwp66 Cwp66 (VF0591)
	<i>cwpV</i>	hemagglutinin/adhesin CwpV (VF0596)
	<i>ebpA</i>	endocarditis and biofilm-associated pilus tip protein EbpA Ebp pili (VF0538)
	<i>fbaB</i>	fibronectin-binding protein SfbII FBPs (VF0243)
	<i>fbpB</i>	Secreted antigen 85-B FbpB (85B Antigen 85 (VF0311)
	<i>fbpC</i>	Secreted antigen 85-C FbpC (85C Antigen 85 (VF0311)
	<i>fhaB</i>	filamentous hemagglutinin/adhesin FHA (VF0025)
	<i>fhaS</i>	hemagglutinin repeat-containing protein FHA-like proteins (VF0691)
	<i>fimA</i>	type 1 fimbrial protein subunit FimA Type 1 fimbriae (VF0102)
		Type-1 fimbrial protein, A chain precursor Type 1 fimbriae (VF0221)
		type 1 fimbrial major subunit FimA Type 1 fimbriae (VF0102)
		type-1 fimbrial protein subunit A Type 1 fimbriae (VF0102)
	<i>fimA gene</i>	major fimbrial subunit Type 1 fimbriae (VF1211)
	<i>fimB</i>	tyrosine recombinase Type I fimbriae (VF0566)
	<i>fimH</i>	type I fimbriae minor fimbrial subunit FimH, adhesin Type 1 fimbriae (VF0102)
	<i>fimQ</i>	unknown Type 1 pili (VF1200)
	<i>fimT</i>	type IV pilus biogenesis protein FimT Type IV pili (VF0155)
	<i>fimV</i>	hypothetical protein TFP (VF1334)
	<i>flpE</i>	Flp pilus assembly protein, ATPase Flp type IV pili (VF0476)
	<i>flpG</i>	Flp pilus assembly protein FlpG Flp type IV pili (VF0476)
	<i>flpH</i>	Flp pilus assembly protein FlpH Flp type IV pili (VF0476)
	<i>flpI</i>	Flp pilus assembly protein FlpI Flp type IV pili (VF0476)
	<i>fnbB</i>	fibronectin binding protein FnbB CAA80122 (VF1079)
	<i>fne</i>	fibronectin-binding protein Fne AAL09480 (VF1080)
	<i>GBS_RS03565</i>	major subunit protein PI-1 (VF0527)
	<i>hhdA</i>	hemagglutinin repeat-containing protein Hemolysin (VF0757)
	<i>hia</i>	adhesin Hia/Hsf (VF0039)
	<i>hia/hsf</i>	Hia/Hsf adhesin N-terminal domain-containing protein Hia/Hsf (VF0039)
	<i>inlF</i>	internalin F InlF (VF0447)
	<i>inlJ</i>	internalin J InlJ (VF0445)
	<i>M28_RS06795</i>	LPXTG cell wall anchor domain-containing protein AspA M28_Spy1325 (VF1071)
	<i>PA4298</i>	hypothetical protein Type IV pili (VF0928)
	<i>PD_RS00130</i>	pilus assembly protein Type IV pili (VF1240)
	<i>PD_RS02610</i>	pilus assembly protein Type IV pili (VF1240)
	<i>pilC</i>	type IV pilus assembly protein PilC Type IV pili (VF0431)
	<i>pilD</i>	A24 family peptidase Type IV pili (VF0458)
	<i>pilF</i>	type IV pilus biogenesis/stability protein PilW TFP (VF1334)
	<i>pilM</i>	type IV pilus inner membrane platform protein PilM Type IV pili (VF0082)
	<i>pilN</i>	PilN domain-containing protein TFP (VF1334)
	<i>pilO</i>	type 4a pilus biogenesis protein PilO TFP (VF1334)
		type IV pilus inner membrane platform protein PilO Type IV pili (VF0155)
	<i>pilQ</i>	type IV pilus biogenesis protein PilQ Type IV pili (VF0401)
	<i>pilW</i>	type IV fimbrial biogenesis and twitching motility protein PilW Type IV pili (VF0075)
	<i>nilX</i>	type 4 fimbrial biogenesis protein PilX Type IV pili (VF0082)

		type IV pilus minor pilin PilX Type IV pili (VF0075)
		VWA domain-containing protein TFP (VF1334)
	<i>pilY1</i>	type IV pilus biogenesis PilY1-related protein Type IV pili (VF0155)
		type 4 fimbrial biogenesis protein PilY1 Type IV pili (VF0082)
	<i>QU43_RS68240</i>	response regulator transcription factor Cable pilus (VF0700)
	<i>QU43_RS68260</i>	cable pilus usher protein CblC Cable pilus (VF0700)
	<i>rcpA</i>	type II and III secretion system protein family protein The tad locus (VF0754)
	<i>rspA</i>	rhusiopathiae surface protein A RspA (VF1214)
	<i>safC</i>	fimbrial biogenesis outer membrane usher protein Saf (VF0951)
	<i>SAG_RS09105</i>	LPXTG cell wall anchor domain-containing protein Agglutinin receptor (VF1044)
	<i>SAK_RS03870</i>	major subunit protein PI-1 (VF0527)
	<i>SEQ_RS04570</i>	Cna B-type domain-containing protein Cne SEQ_0935 (VF1073)
	<i>SEZ_RS01695</i>	thioester-forming surface-anchored protein Fnz Sez_0313 (VF1082)
	<i>sfaA</i>	S-fimbrial major subunit SfaA S fimbriae (VF0222)
	<i>sipA</i>	signal peptidase I PI-2 (VF0530)
	<i>spaD</i>	SpaD-type pili major subunit SpaD SpaD-type pili (VF0726)
	<i>spaE</i>	SpaD-type pili minor subunit SpaE SpaD-type pili (VF0726)
	<i>spaI</i>	SpaH-type pili minorsubunit SpaI SpaH-type pili (VF0727)
	<i>srtG1</i>	sortase, SrtB family PI-2 (VF0530)
	<i>sspA</i>	antigen I/II family LPXTG-anchored adhesin Antigen I/II (Agl/II) family of oral streptococcal adhesins (VF1043)
		LPXTG cell wall anchor domain-containing protein Antigen I/II (Agl/II) family of oral streptococcal adhesins (VF1043)
	<i>sspB</i>	antigen I/II family LPXTG-anchored adhesin Antigen I/II (Agl/II) family of oral streptococcal adhesins (VF1043)
	<i>SSU05_0965</i>	agglutinin receptor Agglutinin receptor (VF1044)
	<i>staG</i>	fimbrial-like adhesin Sta (VF0953)
	<i>steB</i>	fimbrial biogenesis outer membrane usher protein Ste (VF0957)
	<i>stiC</i>	fimbrial outer membrane usher protein Sti (VF0961)
	<i>tapM</i>	type IV pilus assembly protein PilM Tap type IV pili (VF0475)
	<i>tapO</i>	type IV pilus biogenesis protein PilO Tap type IV pili (VF0475)
	<i>tapY1</i>	type IV pilus biogenesis protein Tap type IV pili (VF0475)
	<i>tcfC</i>	fimbrial outer membrane usher protein TcfC Tcf (VF0964)
	<i>ucaA</i>	uroepithelial cell adherence major pilin UcaA Uroepithelial cell adhesin (UCA) (VF1235)
Antimicrobial activity/Competitive advantage (VFC0325)	<i>hcnB</i>	cyanide-forming glycine dehydrogenase subunit HcnB Hydrogen cyanide production (VF0920)
		hydrogen cyanide synthase subunit HcnB Hydrogen cyanide production (VF0920)
Biofilm (VFC0271)	<i>abaR</i>	DNA-binding HTH domain-containing protein Quorum sensing (VF0471)
	<i>algG</i>	outer membrane protein AlgE Alginate (VF0091)
	<i>algK</i>	alginate biosynthesis protein Alg44 Alginate (VF0091)
	<i>algQ</i>	Alginate regulatory protein AlgQ Alginate (VF0091)
	<i>bspR5</i>	autoinducer binding domain-containing protein Quorum-sensing (VF0433)
	<i>cqsA</i>	CAI-1 autoinducer synthase CAI-1 (VF0405)
	<i>csuA</i>	Csu pilus subunit CsuA Csu fimbriae (VF0461)
	<i>csuC</i>	Csu pilus chaperone protein CsuC Csu fimbriae (VF0461)
	<i>flgK</i>	flagellar hook-associated protein FlgK Lateral flagella (VF0625)
	<i>fliF</i>	flagellar M-ring protein FliF Lateral flagella (VF0625)
	<i>mrkH</i>	transcriptional activator Type 3 fimbriae (VF0567)
	<i>mrkI</i>	LuxR family regulatory protein Type 3 fimbriae (VF0567)
	<i>pmlR/bspR1</i>	N-acylhomoserine lactone dependent regulatory protein Quorum-sensing (VF0433)
	<i>VC_RS04620</i>	hemolysin RbmC (VF0617)
	<i>VP_RS22610</i>	OmpA family protein Lateral flagella (VF0625)
	<i>vpsD</i>	exopolysaccharide biosynthesis glycosyltransferase VpsD VPS (VF0615)
Effector delivery system (VFC0086)	<i>aaiA</i>	type VI secretion system contractile sheath small subunit AAI/SCI-II T6SS (VF1121)
	<i>ankA</i>	Anaplasma T4SS effector protein, ankyrin repeat domain-containing protein A Rvh T4SS secreted effectors (VF0654)
	<i>bapC</i>	invasion protein Bsa T3SS (VF0428)
	<i>bepA</i>	Bartonella effector protein BepA VirB/VirD4 type IV secretion system (VF0369)
	<i>bicA</i>	chaperone Bsa T3SS (VF0428)
	<i>BPE043</i>	type IV secretion system effector T4SS secreted effectors (VF0695)
	<i>bscL</i>	Bordetella Bsc type III secretion system protein BscL TTSS (VF0035)
	<i>bscU</i>	Bordetella Bsc type III secretion system protein BscU TTSS (VF0035)

<i>bspF</i>	type IV secretion system effector, GNAT family N-acetyltransferase T4SS secreted effectors (VF0695)
<i>CBU_1268</i>	Coxiella Dot/Icm type IVB secretion system translocated effector T4SS secreted effectors (VF0696)
<i>CBUD_RS05145</i>	DUF3800 domain-containing protein T4SS secreted effectors (VF0696)
<i>CBUG_RS03760</i>	Coxiella Dot/Icm type IVB secretion system translocated effector T4SS secreted effectors (VF0696)
<i>cdiB</i>	ShlB/FhaC/HecB family hemolysin secretion/activation protein Contact-dependent inhibition CDI system (VF1131)
<i>cdsS</i>	Yop proteins translocation protein S TTSS (VF0344)
<i>cesD</i>	chaperone for EspB and EspD TTSS (VF0191)
<i>copN</i>	Type III secretion system effector CopN TTSS (VF0344)
<i>CV_RS09350</i>	ShlB/FhaC/HecB family hemolysin secretion/activation protein ShlA- ShlB (VF1261)
<i>CV_RS09355</i>	hemagglutinin repeat-containing protein ShlA-ShlB (VF1261)
<i>CV_RS12830</i>	EscU/YscU/HrcU family type III secretion system export apparatus switch protein Cpi-1a + Cpi-1 (SPI-1 like) (VF1259)
<i>cyp143</i>	cytochrome P450 143 cyp143 ESX-5 (VF0402)
	cytochrome P450 Cyp143 ESX-5 (VF0402)
	cytochrome P450 ESX-5 (VF0402)
<i>dotU</i>	type VI secretion system protein TssL, short form T6SS-II (VF0783)
<i>dotU/tssL</i>	type VI secretion system protein, DotU/TssL family T6SS (VF0569)
<i>eccA2</i>	type VII secretion system ESX-2 AAA family ATPase EccA2 ESX-2 (T7SS) (VF0834)
	CbxX/CfqX family protein ESX-2 (T7SS) (VF0834)
<i>eccB4</i>	type VII secretion protein EccB ESX-4 (T7SS) (VF0836)
	hypothetical protein ESX-4 (T7SS) (VF0836)
	type VII secretion system ESX-4 subunit EccB4 ESX-4 (T7SS) (VF0836)
<i>eccCb1</i>	ESX-1 type VII secretion system FtsK/SpoIIIE family protein EccCb1 ESX-1 (VF0315)
<i>eccD1</i>	ESX-1 type VII secretion system EccD1 ESX-1 (VF0315)
<i>eccD4</i>	type VII secretion integral membrane protein EccD ESX-4 (T7SS) (VF0836)
	hypothetical protein ESX-4 (T7SS) (VF0836)
<i>epsC</i>	type II secretion system protein GspC Eps T2SS (VF0613)
<i>epsI</i>	type II secretion system minor pseudopilin GspI Eps T2SS (VF0613)
<i>esaA</i>	type VII secretion system protein EsaA Type VII secretion system (VF0403)
<i>ETAE_RS04155</i>	SycD/LcrH family type III secretion system chaperone T3SS (VF1266)
<i>ETAE_RS04325</i>	hemagglutinin repeat-containing protein EthA-EthB (VF1267)
<i>exel</i>	general secretion pathway protein I Exe T2SS (VF0478)
<i>exlA</i>	hemolysin Exolysin (VF0582)
<i>fha1</i>	type VI secretion system forkhead-associated protein Fha1 HSI-1 (VF0334)
<i>gspC</i>	type II secretion system protein GspC Gsp (VF1177)
	general secretion pathway protein C T2SS (VF0333)
<i>gspl</i>	type II secretion system minor pseudopilin GspI Gsp (VF1296)
<i>gspK</i>	type II secretion system minor pseudopilin GspK Gsp (VF1177)
	general secretion pathway protein K T2SS (VF0333)
	general secretion pathway protein K T2SS (VF1335)
<i>hecB</i>	HecB HecA-HecB (VF1273)
<i>hopAC1</i>	membrane-targeted effector domain-containing toxin P. syringae TTSS effectors (VF0911)
<i>hpmA</i>	calcium-independent hemolysin HpmA HpmA-HpmB (VF1292)
<i>hrcR</i>	HrcR T3SS (VF1283)
<i>hrcV</i>	type III secretion system export apparatus subunit SctV P. syringae TTSS (VF0909)
<i>hrpQ</i>	type III secretion system inner membrane ring subunit SctD P. syringae TTSS (VF0909)
<i>hsiE1</i>	type VI secretion system hcp secretion island protein HsiE1, interacting with HsiB1 to form a novel subcomplex of the T6SS HSI-1 (VF0334)
<i>hsiH1/tssG</i>	type VI secretion system hcp secretion island protein HsiH1 HSI-1 (VF0334)
<i>icmO/dotL</i>	Dot/Icm type IV secretion system coupling protein IcmO/DotL Dot/Icm (VF0156)
	Type IVB secretion system protein IcmO/DotL, an inner membrane ATPase of the type IV coupling protein complex Dot/Icm T4SS (VF0503)
<i>impA/tssA</i>	type VI secretion system protein TssA T6SS (VF0569)

<i>ipgC</i>	type III secretion system chaperone for IpaB and IpaC TTSS (VF0118)
<i>KPK_RS21755</i>	A24 family peptidase Pul (VF0785)
<i>legD2</i>	Dot/Icm type IV secretion system effector LegD2 Dot/Icm T4SS secreted effectors (VF0798)
<i>lidL</i>	Dot/Icm type IV secretion system effector LidL Dot/Icm T4SS secreted effectors (VF0798)
<i>lip1</i>	lipoprotein HSI-1 (VF0334)
<i>LPG_RS09565</i>	Dot/Icm type IV secretion system effector Dot/Icm T4SS secreted effectors (VF0798)
<i>lspI</i>	general secretion pathway protein I Lsp T2SS (VF0154)
<i>lvhB10</i>	TrbI/VirB10 family protein Lvh (Legionella vir homologs) type IVA secretion system (VF0799)
<i>mycP1</i>	ESX-1 type VII secretion system subtilisin-like serine protease MycP1 ESX-1 (VF0315)
<i>mycP2</i>	type VII secretion system ESX-2 serine protease mycosin MycP2 ESX-2 (T7SS) (VF0834)
<i>mycP3</i>	Type VII secretion system membrane-anchored mycosin MycP3 ESX-3 (VF0521)
<i>mycP4</i>	type VII secretion-associated serine protease mycosin ESX-4 (T7SS) (VF0836)
	type VII secretion system ESX-4 serine protease mycosin MycP4 ESX-4 (T7SS) (VF0836)
	serine protease ESX-4 (T7SS) (VF0836)
	Putative protease ESX-4 (T7SS) (VF0836)
<i>NGR_RS11030</i>	hypothetical protein T3SS2 (VF1299)
<i>PA0680</i>	HxcV pseudopilin Hxc (VF0945)
<i>PAU_RS17190</i>	ATPase T3SS-2 (VF1287)
<i>pldA/tle5a</i>	phospholipase D HSI-2 T6SS secreted effectors (VF0915)
<i>ravL</i>	Dot/Icm type IV secretion system effector RavL Dot/Icm T4SS secreted effectors (VF0798)
<i>ROD_RS26480</i>	SycD/LcrH family type III secretion system chaperone T3SS (VF1262)
<i>RS_RS15635</i>	type II secretion system protein N Gsp (VF1296)
<i>rvhB4b</i>	type IV secretion system protein VirB4 precursor Rvh T4SS (VF0488)
<i>rvhB6c</i>	VirB6-like protein of the type IV secretion system Rvh T4SS (VF0488)
<i>sctD</i>	type III secretion system inner membrane ring subunit SctD T3SS (VF1270)
<i>sctJ</i>	type III secretion inner membrane ring lipoprotein SctJ T3SS (VF1266)
<i>sctR</i>	type III secretion system export apparatus subunit SctR TTSS (chromosomally encoded) (VF1023)
<i>sctU</i>	type III secretion system export apparatus subunit SctU TTS1 (VF0702)
<i>shlA</i>	hemolysin ShlA-ShlB (VF1297)
<i>sicA</i>	type III secretion system translocator chaperone SicA Cpi-1a + Cpi-1 (SPI-1 like) (VF1259)
	chaperone for SipC and SipB TTSS (SPI-1 encode) (VF0116)
<i>SPRO_RS22995</i>	hemagglutinin repeat-containing protein ShlA-ShlB (VF1298)
<i>ssaS</i>	type III secretion system minor export apparatus protein SsaS TTSS (SPI-2 encode) (VF0321)
<i>STM0280</i>	type VI secretion system-associated lipoprotein SCI (Salmonella centrisome island) (VF0974)
<i>tagS</i>	type VI secretion associated protein TagS, forming a stable inner membrane complex with TagT HSI-1 (VF0334)
<i>tke5</i>	putative T6SS effector Putida-T6SS secreted effectors (VF1354)
<i>tke7</i>	putative T6SS effector Putida-T6SS secreted effectors (VF1354)
<i>trwE</i>	Trw type IV secretion system VirB10 homolog, trwE protein Trw type IV secretion system (VF0372)
<i>trwG</i>	Trw type IV secretion system VirB8 homolog, trwG protein Trw type IV secretion system (VF0372)
<i>trwI1</i>	Trw type IV secretion system trwI1 protein required for pilus elongation Trw type IV secretion system (VF0372)
<i>trwI2</i>	Trw type IV secretion system trwI2 protein required for pilus elongation Trw type IV secretion system (VF0372)
<i>trwI3</i>	type IV secretion system protein Trw type IV secretion system (VF0372)
<i>trwI4</i>	type IV secretion system protein Trw type IV secretion system (VF0372)
<i>tssF</i>	type VI secretion system baseplate subunit TssF HSI-2 (VF0943)
<i>tssF-5</i>	Type VI secretion system protein TssF T6SS-1 (VF0429)
<i>tssJ</i>	type VI secretion system lipoprotein TssJ HSI-2 (VF0943)
<i>tssK3</i>	type VI secretion system baseplate subunit TssK Putida K3-T6SS (VF1353)
<i>tssL</i>	type VI secretion system protein TssL, long form Imp (VF1258)

	<i>vbhB8</i>	virB8 family protein VirB-homolog (vbh) type IV secretion system (VF0680)
	<i>VC0395_RS15460</i>	ExeA family protein Eps T2SS (VF0613)
	<i>vceC</i>	type IV secretion system effector VceC T4SS secreted effectors (VF0695)
	<i>vcrD2</i>	type III secretion system protein T3SS2 (VF0409)
	<i>vgrG-5</i>	Type VI secretion system protein VgrG5 T6SS-1 secreted effectors (VF1192)
		Type VI secretion system protein VgrG5 T6SS-1 (VF0429)
	<i>virB6</i>	type IV secretion system channel protein VirB6 VirB type IV secretion system (VF0365)
	<i>virB8</i>	type IV secretion system virB8 protein homolog VirB/VirD4 type IV secretion system (VF0369)
	<i>virB8-1</i>	type IV secretion protein VirB8 Rvh T4SS (VF0487)
	<i>virB9</i>	P-type conjugative transfer protein VirB9 T4SS (VF1269)
	<i>virB9-2</i>	type IV secretion system protein VirB9 Rvh T4SS (VF0487)
	<i>vpdC</i>	Dot/Icm type IV secretion system effector VpdC Dot/Icm T4SS secreted effectors (VF0798)
	<i>vscJ</i>	type III secretion system IM ring protein VscJ T3SS1 (VF0408)
	<i>vscS</i>	type III secretion system C-ring protein VscS T3SS1 (VF0408)
	<i>XBJ1_RS01135</i>	hemagglutinin repeat-containing protein XhIA-XhIB (VF1305)
	<i>XCC_RS03435</i>	prepilin-type N-terminal cleavage/methylation domain-containing protein Xps (VF1304)
	<i>xcpX</i>	general secretion pathway protein K Xcp secretion system (VF0084)
	<i>XNC1_RS19230</i>	hemagglutinin repeat-containing protein XhIA-XhIB (VF1306)
	<i>yadA</i>	trimeric autotransporter adhesin YadA YadA (VF0133)
	<i>YE_RS12055</i>	hemagglutinin repeat-containing protein Y2267-y2268 (VF1034)
	<i>YE105_RS01820</i>	EscV/YscV/HrcV family type III secretion system export apparatus protein TTSS (chromosomally encoded) (VF1023)
	<i>YE105_RS01850</i>	EscS/YscS/HrcS family type III secretion system export apparatus protein TTSS (chromosomally encoded) (VF1023)
	<i>ygeG</i>	SycD-like chaperone family TPR-repeat-containing protein ETT2 (VF1161)
	<i>yhlA</i>	YhlA YhlA-YhlB (VF1032)
	<i>YPO_RS02340</i>	EscV/YscV/HrcV family type III secretion system export apparatus protein TTSS (chromosomally encoded) (VF1023)
	<i>YPSIP31758_RS01000</i>	type IV secretory system conjugative DNA transfer family protein lcm/dot type IVB locus (VF1026)
	<i>YPSIP31758_RS02865</i>	hemagglutinin repeat-containing protein ShIA-ShIB (VF1035)
	<i>ysaC</i>	type III secretion system outer membrane ring subunit SctC Ysa TTSS (VF1025)
	<i>yst1I</i>	type II secretion system minor pseudopilin Gspl Yst1 T2SS (VF0635)
	<i>yst1O</i>	A24 family peptidase Yst1 T2SS (VF0635)
Exoenzyme (VFC0251)	<i>cbpD</i>	choline binding protein D Autolysin (VF0143)
	<i>cppA</i>	CppA family protein C3-degrading protease (VF1054)
	<i>PA4541</i>	hypothetical protein LepA (VF0946)
	<i>sdn</i>	DNA/RNA non-specific endonuclease Streptodornase (VF1068)
	<i>smcL</i>	sphingomyelinase-c SMase (VF0061)
Exotoxin (VFC0235)	<i>acpC</i>	acyl carrier protein AcpC -haemolysin/cytolysin (VF0279)
	<i>alo</i>	thiol-activated cytolysin ALO (VF0534)
	<i>BCB4264_RS25790</i>	anthrolysin O/cereolysin O family cholesterol-dependent cytolysin ALO (VF0534)
	<i>BCERKBAB4_RS24660</i>	anthrolysin O/cereolysin O family cholesterol-dependent cytolysin ALO (VF0534)
	<i>cdtB</i>	cytolethal distending toxin B CDT (VF0115)
	<i>cfa3</i>	beta-ketoacyl-[acyl-carrier-protein] synthase family protein Phytotoxin coronatine (VF0916)
	<i>hIb</i>	beta-hemolysin -hemolysin (VF0002)
	<i>pat1</i>	patatin-like phospholipase Phospholipase A2 (VF0494)
	<i>plcC</i>	phospholipase C Phospholipase C (VF0320)
	<i>PSPTO_RS24365</i>	alanyl-tRNA editing protein Phytotoxin coronatine (VF0916)
	<i>rtxA</i>	structural toxin protein RtxA RtxA (VF0259)
		multifunctional autoprocessing repeats-in-toxins (MARTX RtxA (VF0482)
		RTX toxin RtxA MARTX (VF0265)
	<i>spa</i>	Immunoglobulin G binding protein A precursor SpA (VF0017)
	<i>sph</i>	sphingomyelin phosphodiesterase Sphingomyelinase (SMase) (VF0663)
	<i>VIBHAR_RS18025</i>	SGNH/GDSL hydrolase family protein TLH (VF0610)
Immune modulation (VFC0258)]	<i>aepY</i>	phosphonopyruvate decarboxylase Capsule (VF0323)
	<i>aftB</i>	terminal beta-(1->2)-arabinofuranosyltransferase LAM (VF0318)
	<i>aftD</i>	alpha-(1->3)-arabinofuranosyltransferase LAM (VF0318)
	<i>atf</i>	acyltransferase GPL locus (VF0841)
	<i>bexB'</i>	BexB' Capsule (VF0043)
	<i>bplE</i>	probable glycosyl transferase LPS (VF0033)

<i>brkB</i>	serum resistance protein Brk (VF0034)
<i>BT9727_RS25825</i>	glycosyltransferase family 4 protein Polysaccharide capsule (VF0659)
<i>C8J_RS05635</i>	capsular polysaccharide synthesis protein LOS (VF0326)
<i>capB</i>	capsule biosynthesis protein CapB Capsule (VF0543)
<i>CFF8240_RS06855</i>	polysaccharide deacetylase family protein LOS (VF0326)
<i>CGSHIEE_RS06420</i>	sulfatase-like hydrolase/transferase LOS (VF0044)
<i>CGSHIGG_RS08225</i>	sulfatase-like hydrolase/transferase LOS (VF0044)
<i>chp</i>	chemotaxis-inhibiting protein CHIPS CHIPS (VF0424)
<i>Cj1137c</i>	glycosyltransferase LOS (VF0326)
<i>Cj1418c</i>	hypothetical protein Capsule (VF0323)
<i>CJE_RS08020</i>	glycosyltransferase Capsule (VF0323)
<i>CJJ81176_RS06820</i>	hypothetical protein Capsule (VF0323)
<i>cpsA</i>	CpsA responsible for transfer of CPS from the membrane lipid to the cell wall peptidoglycan Capsule (VF0274)
<i>ctrC</i>	capsule polysaccharide export inner-membrane protein CtrC Capsule (VF0079)
<i>drnC</i>	ABC transporter permease PDIM (VF0309)
<i>F7308_RS05550</i>	glycosyl transferase Capsule (VF0543)
<i>futA</i>	fucosyltransferase Lewis antigen (VF0057)
<i>GBS_RS06605</i>	tyrosine-protein phosphatase CpsB Capsule (VF0274)
<i>GBS_RS06610</i>	LCP family protein Capsule (VF0274)
<i>gtf3</i>	glycosyltransferase GPL locus (VF0841)
<i>hasA</i>	hyaluronan synthase HasA Hyaluronic acid capsule (VF0244)
	hyaluronan synthase HasA Capsule (VF0144)
<i>HD_RS06500</i>	phosphoethanolamine transferase LOS (VF0044)
<i>htrB</i>	lipid A biosynthesis lauroyl acyltransferase LPS (VF0370)
<i>JJD26997_RS08110</i>	hypothetical protein Capsule (VF0323)
<i>KPK_RS08095</i>	N-acetyl-alpha-D-glucosaminyl-diphospho-ditrans, octakis-undecaprenol 3-alpha-mannosyltransferase WbdC LPS (VF0561)
<i>KPN_RS13390</i>	glycosyltransferase LPS (VF0561)
<i>KPN_RS13485</i>	capsular polysaccharide synthesis protein Capsule (VF0560)
<i>KPR_RS08980</i>	hypothetical protein Capsule (VF0560)
<i>KPR_RS28090</i>	glycosyltransferase family 4 protein Capsule (VF0560)
<i>lgtG</i>	glycosyltransferase LOS (VF0078)
<i>licC</i>	phosphocholine cytidyltransferase LOS (VF0044)
<i>lpsA</i>	glycosyltransferase LOS (VF0044)
<i>lpt6</i>	phosphoethanolamine (Petr) LOS (VF0044)
<i>lpxB</i>	lipid-A-disaccharide synthase LPS (VF0370)
<i>lsgF</i>	glycosyltransferase LOS (VF0044)
<i>M3Q_RS01460</i>	MaoC family dehydratase Capsule (VF0465)
<i>manB</i>	phosphomannomutase LPS (VF0542)
<i>mce2C</i>	MCE family protein Mce2 (VF0825)
	virulence factor Mce family protein Mce2 (VF0825)
<i>mce3C</i>	MCE family protein Mce3 (VF0826)
	Mce family protein Mce3C Mce3 (VF0826)
	virulence factor Mce family protein Mce3 (VF0826)
<i>mce4B</i>	Hypothetical MCE-family protein Mce4 (VF0827)
<i>mce4C</i>	virulence factor Mce family protein Mce4 (VF0827)
<i>N559_RS09035</i>	glycosyltransferase Capsule (VF0560)
<i>napA</i>	neutrophil activating protein NapA HP-NAP (VF0052)
<i>neuA</i>	N-acetylneuraminate cytidyltransferase Capsule (VF0274)
<i>NTHI_RS05825</i>	sulfatase-like hydrolase/transferase LOS (VF0044)
<i>opsX/rfaC</i>	heptosyltransferase I LOS (VF0044)
<i>PA14_RS09490</i>	heparinase II/III family protein LPS (VF0085)
<i>PA14_RS29915</i>	glycosyltransferase family 4 protein LPS (VF0085)
	acyltransferase PDIM (VF0309)
<i>papA5</i>	phthiocerol/phthiodiolone dimycocerosyl transferase PDIM (VF0309)
<i>pmm</i>	phosphomannomutase LPS (VF0367)
<i>pseG</i>	UDP-2,4-diacetamido-2,4, 6-trideoxy-beta-L-altropyranose hydrolase Capsule (VF0465)
<i>pseH</i>	UDP-4-amino-4, 6-dideoxy-N-acetyl-beta-L-altrosamine N-acetyltransferase Capsule (VF0465)
<i>rhlA</i>	rhamnosyltransferase chain A Rhamnolipid (VF0089)
<i>rmt3</i>	Methyltransferase MtfD GPL locus (VF0841)
	class I SAM-dependent methyltransferase GPL locus (VF0841)
<i>SAK_RS06325</i>	tyrosine-protein phosphatase CpsB Capsule (VF0274)
<i>SAK_RS06330</i>	LCP family protein Capsule (VF0274)
<i>SGO_RS09910</i>	capsular polysaccharide biosynthesis protein CpsC Capsule (VF0144)
<i>SGO_RS09915</i>	tyrosine protein phosphatase Capsule (VF0144)
<i>siaC/synC</i>	polysialic acid capsule biosynthesis protein SiaC Capsule (VF0079)
<i>SP_RS01735</i>	polysaccharide pyruvyl transferase family protein Capsule (VF0144)

	<i>SPD_RS01750</i>	glycosyltransferase family 4 protein Capsule (VF0144)
	<i>SPH_RS02240</i>	LicD family protein Capsule (VF0144)
	<i>SSA_RS07395</i>	glycosyltransferase family 4 protein Capsule (VF0144)
	<i>SSU05_0567</i>	Cps2D Capsule (VF0144)
	<i>STER_RS05315</i>	tyrosine protein phosphatase Capsule (VF0144)
	<i>STR_RS05260</i>	tyrosine protein phosphatase Capsule (VF0144)
	<i>STR_RS05265</i>	LCP family protein Capsule (VF0144)
	<i>STU_RS14645</i>	tyrosine protein phosphatase Capsule (VF0144)
	<i>STU_RS14650</i>	LCP family protein Capsule (VF0144)
	<i>waaF</i>	lipopolysaccharide heptosyltransferase II LOS (VF0326)
	<i>wbmJ</i>	glycosyltransferase LPS (VF0033)
	<i>wbmO</i>	glycosyltransferase family 1 protein LPS (VF0033)
	<i>wbpZ</i>	mannosyltransferase C LPS (VF0367)
	<i>wzm</i>	capsular polysaccharide export ABC transporter transmembrane protein Capsule I (VF0436)
	<i>YE_RS15465</i>	flippase O-antigen (VF0392)
	<i>yhbX</i>	phosphoethanolamine transferase LOS (VF0044)
	<i>YPK_RS15965</i>	MATE family efflux transporter O-antigen (VF0392)
	<i>YPO_RS16495</i>	NAD(P)/FAD-dependent oxidoreductase O-antigen (VF0392)
	<i>YPTB_RS05535</i>	NAD(P)-binding protein O-antigen (VF0392)
	<i>zmpC</i>	immunoglobulin A1 protease IgA1 protease (VF0147)
Invasion (VFC0083)	<i>ctad1</i>	Chlamydia SH3 domain containing adhesin and invasin Ctad1 Ctad1 (VF1348)
	<i>inlA</i>	Internalin A InlA (VF0069)
	<i>kpsD</i>	Polysialic acid transport protein precursor K1 capsule (VF0239)
Motility (VFC0204)	<i>BCE_RS08885</i>	flagellar biosynthetic protein FliQ Peritrichous flagella (VF0670)
	<i>CD630_RS01740</i>	class I SAM-dependent methyltransferase Peritrichous flagella (VF0732)
	<i>cheY1</i>	chemotaxis two-component response regulator CheY1 Flagella (VF0430)
	<i>cheZ</i>	protein phosphatase CheZ Peritrichous flagella (VF0967)
		protein phosphatase CheZ Peritrichous flagella (VF1154)
		chemotaxis regulator CheZ Flagella (VF0430)
	<i>flaA</i>	flagellin Flagella (VF0114)
	<i>flaJ</i>	flagellar specific chaperone FlaJ/FliS Polar flagella (VF0473)
	<i>flgB</i>	flagellar basal body rod protein FlgB Peritrichous flagella (VF0732)
		flagellar basal body rod protein FlgB Flagella (VF0273)
		flagellar basal body rod protein FlgB Peritrichous flagella (VF1154)
		flagellar basal body rod protein FlgB Peritrichous flagella (VF0967)
		flagellar basal body rod protein FlgB Flagella (VF0157)
	<i>flgD</i>	flagellar basal-body rod modification protein FlgD Flagella (VF0051)
	<i>flgE</i>	flagellar hook protein FlgE Flagella (VF0114)
	<i>flgE_1</i>	flagellar hook protein Flagella (VF0051)
	<i>flgH</i>	flagellar L-ring protein precursor FlgH Flagella (VF0114)
		flagellar L-ring protein precursor FlgH Flagella (VF0051)
	<i>flgK</i>	flagellar hook-associated protein FlgK Flagella (VF0685)
	<i>flgL</i>	flagellar hook-associated protein 3 FlgL Flagella (VF0273)
	<i>flgN</i>	flagella synthesis protein FlgN Flagella (VF0273)
	<i>flgP</i>	Vibrio-specific flagellar H-ring component FlgP Flagella (VF0519)
	<i>flhB</i>	flagellar biosynthesis protein Flagella (VF0051)
		flagellar biosynthetic protein FlhB Flagella (VF0373)
	<i>flhF</i>	flagellar biosynthesis protein Flagella (VF0051)
	<i>flhG</i>	flagellar biosynthesis protein FlhG Flagella (VF0430)
	<i>fliF</i>	flagellar M-ring protein FliF Flagella (VF0114)
		flagellar MS-ring protein FliF Peritrichous flagella (VF0796)
		flagellar M-ring protein FliF Flagella (VF0519)
		flagellar M-ring protein FliF Flagella (VF0051)
	<i>fliH</i>	flagellar assembly protein H Flagella (VF0273)
	<i>fliK</i>	flagellar hook-length control protein FliK Flagella (VF0273)
	<i>fliL</i>	flagellar basal body protein FliL Flagella (VF0273)
	<i>fliO</i>	flagellar biosynthesis protein, FliO Polar flagella (VF0473)
		flagellar protein FliO Flagella (VF0273)
	<i>fliQ</i>	flagellar biosynthesis protein FliQ Peritrichous flagella (VF0796)
	<i>fliR</i>	flagellar biosynthesis protein FliR Polar flagella (VF0473)
		flagellar biosynthesis protein Flagella (VF0519)
	<i>fliS</i>	flagellar protein FliS Flagella (VF0519)
		flagellar protein FliS Flagella (VF0430)
	<i>fliM</i>	FliM protein Polar flagella (VF0473)
	<i>lafA</i>	lateral flagellin LafA Lateral flagella (VF0474)
	<i>lafU</i>	LafU, chemotaxis protein Lateral flagella (VF0474)
	<i>lfiN</i>	lateral flagellar motor switch protein, lfiN Lateral flagella (VF0474)
	<i>lfiQ</i>	lateral flagellar biosynthetic protein, LfiQ Lateral flagella (VF0474)

	<i>lmo0685</i>	flagellar motor protein MotA Peritrichous flagella (VF0796)
	<i>LPG_RS13035</i>	flagellar protein Flagella (VF0157)
	<i>motA</i>	flagellar motor stator protein MotA Peritrichous flagella (VF0670)
	<i>pomB2</i>	flagellar motor protein MotB Polar flagella (VF0473)
Nutritional/Metabolic factor (VFC0272)	<i>ptmB</i>	acylneuraminate cytidyltransferase Flagella (VF0114)
		acylneuraminate cytidyltransferase O-linked flagellar glycosylation (VF0720)
	<i>Cj0178</i>	TonB-dependent outer membrane receptor Transferrin and Lactoferrin-mediated Iron uptake (VF0725)
	<i>acsB</i>	4-hydroxy-2-oxovalerate aldolase Achromobactin biosynthesis and transport (VF0922)
	<i>ansP</i>	major facilitator transporter AnsP (VF0554)
	<i>ASA_RS16550</i>	cysteine hydrolase Heme uptake system (VF0650)
	<i>DDA3937_RS07715</i>	TonB-dependent receptor Achromobactin (VF1243)
	<i>fadE28</i>	Probable acyl-CoA dehydrogenase FadE FadE28 (VF0844)
		acyl-CoA/acyl-ACP dehydrogenase FadE28 (VF0844)
	<i>hasA</i>	unknown protein HasA (VF1254)
	<i>hasAp</i>	heme acquisition protein HasA HasA-type hemophore-mediated heme uptake system (VF0939)
	<i>hasR</i>	unknown protein HasA (VF1254)
	<i>hbpC</i>	porin family protein Heme binding proteins (VF0678)
	<i>hutA</i>	TonB-dependent hemoglobin/transferrin/lactoferrin family receptor Heme receptor HutA (VF0629)
	<i>ileP</i>	major facilitator transporter IleP (VF0556)
	<i>ilsA</i>	NEAT domain-containing leucine-rich repeat protein IlsA (VF0585)
	<i>iroB</i>	putative glycosyl transferase Salmochelin (VF0972)
		glucosyltransferase IroB Salmochelin siderophore (VF0230)
		glucosyltransferase IroB Sal (VF0563)
	<i>iroC</i>	salmochelin biosynthesis C-glycosyltransferase IroB Salmochelin synthesis and transport (VF0983)
	<i>iutA</i>	ferric aerobactin receptor precursor IutA Aerobactin (VF0229)
		ferric aerobactin receptor IutA Aerobactin (VF0565)
	<i>mbtI</i>	Isochorismate synthase MbtI Mycobactin (VF0299)
	<i>mbtJ</i>	Putative acetyl hydrolase MbtJ Mycobactin (VF0299)
	<i>mbtL</i>	Acyl carrier protein (ACP) Mycobactin (VF0299)
	<i>mbtN</i>	Acyl-CoA dehydrogenase MbtN Mycobactin (VF0299)
	<i>mce1C</i>	MCE family protein Mce1 (VF0824)
	<i>mce1D</i>	virulence factor Mce family protein Mce1 (VF0824)
		MCE-family protein MCE1D Mce1 (VF0824)
		Mce family protein Mce1D Mce1 (VF0824)
	<i>PA3404</i>	hypothetical protein HasA-type hemophore-mediated heme uptake system (VF0939)
	<i>PA4705</i>	hypothetical protein Direct heme uptake system (VF0941)
	<i>pdtorfN</i>	putative ABC transporter Pyridine-2,6-dithiocarboxylic acid (PDTC) (VF0938)
	<i>PFL_RS27285</i>	heme acquisition protein HasA HasA-type hemophore-mediated heme uptake system (VF0940)
	<i>pvcA</i>	paerucumarin biosynthesis protein PvcA Pyoverdine (VF0094)
	<i>pvdF</i>	pyoverdine synthetase F Pyoverdine (VF0094)
	<i>pvdP</i>	tyrosinase required for pyoverdine maturation Pyoverdine (VF0094)
	<i>QU43_RS45025</i>	malleobactin biosynthesis enzyme MbaF Ornibactin (VF0701)
	<i>srtB</i>	class B sortase NEAT-type hemophore-mediated heme uptake system (VF0672)
	<i>vctA</i>	TonB-dependent siderophore receptor Enterobactin receptor VctA (VF0639)
	<i>vctP</i>	siderophore ABC transporter substrate-binding protein VctPDGC system (VF0638)
Post-translational modification (VFC0315)	<i>pglD</i>	UDP-N-acetylglucosamine N-acetyltransferase N-linked protein glycosylation (VF0721)
	<i>pglH</i>	glycosyltransferase N-linked protein glycosylation (VF0721)
	<i>srtB</i>	class B sortase Sortase B (VF0789)
		hypothetical protein Sortase B (VF0789)
Stress survival (VFC0282)	<i>acpC</i>	acid phosphatase Acid phosphatase (VF0547)
Others (VFC0346)	<i>fliQ</i>	flagellar biosynthesis protein FliQ Lateral flagella (VF1028)
	<i>lidL</i>	Dot/Icm type IV secretion system effector LidL Enh loci (VF0163)
	<i>mce7A</i>	MCE family protein Mce7 (VF0830)
	<i>mce7E</i>	
	<i>mce8B</i>	
	<i>mce8E</i>	MCE family protein Mce8 (VF0831)
	<i>msbB2</i>	
	<i>vapA5</i>	HigA family addition module antidote protein Virulence-associated proteins (VF0679)
	<i>YPO_RS04505</i>	flagellar biosynthetic protein FliR Lateral flagella (VF1028)

Annex 3. Unique virulence factors in the biochemical pregnancy (BP) group.

Category	VF gene name	Description
Adherence (VFC0001)	<i>CGSHIEE_RS03620</i>	filamentous hemagglutinin N-terminal domain-containing protein HMW1/HMW2 (VF0038)
	<i>CGSHIEE_RS05350</i>	filamentous hemagglutinin N-terminal domain-containing protein HMW1/HMW2 (VF0038)
	<i>CPE_RS09465</i>	type II secretion system F family protein Type IV pili (VF0731)
	<i>fss2</i>	Enterococcus faecalis surface protein Fss2, fibrinogen binding protein Fibrinogen binding protein (VF0751)
	<i>hcpA</i>	prepilin peptidase-dependent pilin Hemorrhagic E.coli pilus (HCP) (VF1141)
	<i>hmw1A</i>	adhesin HMW1A, two-partner secretion (TPS HMW1/HMW2 (VF0038)
	<i>hmw2A</i>	adhesin HMW2A, two-partner secretion (TPS HMW1/HMW2 (VF0038)
	<i>MCR_RS01830</i>	pilin Type IV pili (VF1228)
	<i>mshM</i>	MSHA biogenesis protein MshM MSHA pili (VF0515)
	<i>NTHI_RS06855</i>	filamentous hemagglutinin N-terminal domain-containing protein HMW1/HMW2 (VF0038)
	<i>NTHI_RS09145</i>	filamentous hemagglutinin N-terminal domain-containing protein HMW1/HMW2 (VF0038)
	<i>pilA</i>	prepilin peptidase-dependent pilin Type IV pili (VF0753)
		type IV fimbrial pilin protein Type IV pili (VF0431)
		type 4 fimbrial precursor PilA Type IV pili (VF0082)
	<i>pilE</i>	type 4 fimbrial biogenesis protein PilE Type IV pili (VF0082)
	<i>pilE/pilA</i>	Type IV pilus subunit Type IV pili (VF0458)
	<i>pilE5</i>	type IV pilin protein Type IV pili (VF0458)
		prepilin-type N-terminal cleavage/methylation domain-containing protein Type IV pili (VF0458)
	<i>pilO</i>	type IV pilus inner membrane platform protein PilO Type IV pili (VF0401)
	<i>pilP</i>	type IV pilus biogenesis protein PilP Type IV pili (VF0082)
	<i>pilV</i>	type IV pilus biogenesis protein PilV Type IV pili (VF0082)
	<i>pilX</i>	type IV pilus minor pilin PilX Type IV pili (VF0075)
	<i>sapA</i>	DUF11 domain-containing protein Surface-anchored pilus proteins (VF0736)
	<i>tapA</i>	type IV pilus pilin subunit, pilA-like Tap type IV pili (VF0475)
	<i>tapM</i>	type IV pilus assembly protein PilM Tap type IV pili (VF0475)
	<i>tapV</i>	type IV pilus assembly protein TapV Tap type IV pili (VF0475)
Biofilm (VFC0271)	<i>cupA4</i>	fimbrial subunit CupA4 CupA fimbriae (VF0924)
	<i>mucA</i>	alkaline metalloproteinase precursor Alginate (VF0091)
	<i>mucB</i>	anti-sigma factor MucA, inhibitor of alg gene expression Alginate (VF0091)
	<i>PA14_RS24370</i>	fimbrial protein CupD fimbriae (VF0927)
Effector delivery system (VFC0086)	<i>APECO1_RS01240</i>	DUF4150 domain-containing protein ACE T6SS (VF1122)
	<i>atsK</i>	Type VI secretion system protein T6SS (VF0480)
	<i>bapB</i>	acyl carrier protein Bsa T3SS (VF0428)
	<i>bopN</i>	Bordetella Bsc type III secretion system secreted protein BopN TTSS secreted effectors (VF0687)
	<i>BPE005</i>	type IV secretion system effector, cyclic nucleotide-binding domain-containing protein T4SS secreted effectors (VF0695)
	<i>cdsC</i>	Type III secretion system protein CdsC TTSS (VF0344)
	<i>dotU</i>	Type VI secretion system protein DotU T6SS (VF0480)
	<i>EAMY_RS21060</i>	EscV/YscV/HrcV family type III secretion system export apparatus protein PAI-2 encoded (VF1271)
	<i>ECP_RS01230</i>	DUF4150 domain-containing protein ACE T6SS (VF1122)
	<i>ECs_3718</i>	type III secretion protein EprI ETT2 (VF1161)
	<i>eivA</i>	type III secretion protein EivA ETT2 (VF1161)
	<i>esaD</i>	type VII secretion system secreted protein, a nuclease toxin EsaD Type VII secretion system (VF0403)
	<i>exeN</i>	general secretion pathway protein N Exe T2SS (VF0478)
	<i>gspL</i>	general secretion pathway protein L T2SS (VF1335)
	<i>hcp1</i>	type VI secretion system substrate Hcp1 HSI-1 (VF0334)
	<i>hsiF1/tssE</i>	type VI secretion system hcp secretion island protein HsiF1, a gp25-like protein, but not exhibit lysozyme activity HSI-1 (VF0334)
	<i>icmF</i>	type VI secretion protein VasK T6SS-III (VF0784)
	<i>icmH/tssL</i>	type IVB secretion system protein IcmH/DotU HSI-2 (VF0943)
	<i>icmJ/dotN</i>	type IVB secretion system protein IcmJ/DotN, inner membrane/cytoplasmic protein of the type IV coupling protein complex Dot/Icm T4SS (VF0503)
		Dot/Icm type IV secretion system protein IcmJ/DotN Dot/Icm (VF0156)
	<i>iglB/tssC</i>	Type VI secretion system sheath protein, VipB homolog T6SS (VF0544)
	<i>IS_RS26495</i>	HAD hydrolase-like protein T3SS (VF1301)

	<i>LPG_RS06280</i>	hypothetical protein Lvh (Legionella vir homologs) type IVA secretion system (VF0799)
	<i>LPG_RS11920</i>	Dot/Icm type IV secretion system effector Dot/Icm T4SS secreted effectors (VF0798)
	<i>lvhD4</i>	type IV secretory system conjugative DNA transfer family protein Lvh (Legionella vir homologs) type IVA secretion system (VF0799)
	<i>mxhH</i>	type III secretion system needle filament protein MxiH TTSS (VF0118)
	<i>NUE</i>	hypothetical protein TTSS secreted effectors (VF0711)
		SET domain-containing protein-lysine N-methyltransferase TTSS secreted effectors (VF0711)
	<i>pcrD</i>	type III secretion system protein PcrD TTSS (VF0083)
	<i>RS_RS21180</i>	AWR family protein T3SS (VF1295)
	<i>rvhB6a</i>	VirB6-like protein of the type IV secretion system Rvh T4SS (VF0488)
	<i>sctR</i>	type III secretion system export apparatus subunit SctR T3SS (VF1301)
		type III secretion system export apparatus subunit SctR T3SS (VF1295)
		type III secretion system export apparatus subunit SctR T3SS2 (VF1299)
	<i>tssK</i>	type VI secretion system baseplate subunit TssK EVP (E. tarda virulent protein) (VF1268)
	<i>tssK-5</i>	Type VI secretion system protein TssK T6SS-1 (VF0429)
	<i>UTI89_RS01245</i>	DUF4150 domain-containing protein ACE T6SS (VF1122)
	<i>vasE</i>	type VI secretion system protein VasE VAS T6SS (VF0335)
	<i>vgrG</i>	type VI secretion system tip protein VgrG AAI/SCI-II T6SS (VF1121)
	<i>vgrG1</i>	type VI secretion system tip protein VgrG Putida K1-T6SS (VF1351)
	<i>virB4</i>	Type IV secretion system protein VirB4 Type IV secretion system (VF0325)
	<i>YE105_RS01860</i>	EscU/YscU/HrcU family type III secretion system export apparatus switch protein TTSS (chromosomally encoded) (VF1023)
Exoenzyme (VFC0251)	<i>colA</i>	collagenase Kappa-toxin (VF0388)
	<i>gtfB</i>	glucosyltransferase GtfB Streptococcal glucosyltransferases (VF1051)
	<i>gtfC</i>	glucosyltransferase GtfC Streptococcal glucosyltransferases (VF1051)
	<i>gtfD</i>	glucosyltransferase-S Streptococcal glucosyltransferases (VF1051)
	<i>nanB</i>	neuraminidase B Neuraminidase (VF0148)
	<i>nuc</i>	thermonuclease family protein Thermonuclease nuc (VF0993)
thermonuclease family protein Nuclease (VF0880)		
Exotoxin (VFC0235)	<i>ACR43758</i>	cry4 delta-toxin-like protein Insecticidal crystalline toxins (VF0658)
	<i>hlyA</i>	Hemolysin A -Hemolysin (VF0225)
	<i>PSPPH_RS20435</i>	RHS repeat protein TccC-type insecticidal toxins (VF0921)
	<i>PSPTO_RS22515</i>	RHS repeat protein TccC-type insecticidal toxins (VF0921)
	<i>ymt</i>	murine toxin Ymt (VF0137)
Immune modulation (VFC0258)	<i>A1S_0050</i>	protein tyrosine phosphatase Capsule (VF0465)
	<i>A225_RS19000</i>	ABC transporter permease LPS (VF0561)
	<i>A225_RS19080</i>	glycosyltransferase Capsule (VF0560)
	<i>ABK1_RS00490</i>	oligosaccharide flippase family protein Capsule (VF0465)
	<i>ABZJ_RS06265</i>	oligosaccharide flippase family protein Capsule (VF0465)
	<i>ACICU_RS00465</i>	hypothetical protein Capsule (VF0465)
	<i>BALH_RS26230</i>	lipopolysaccharide biosynthesis protein Polysaccharide capsule (VF0659)
	<i>BCAH187_RS26570</i>	capsular polysaccharide biosynthesis protein Polysaccharide capsule (VF0659)
	<i>BJAB07104_RS00500</i>	glycosyltransferase family 2 protein Capsule (VF0465)
	<i>btpB</i>	Tir domain containing protein BtpB BtpB (VF0522)
	<i>CGSHIEE_RS07300</i>	lipopolysaccharide biosynthesis protein LOS (VF0044)
	<i>CJE_RS07965</i>	NTP transferase domain-containing protein Capsule (VF0323)
	<i>futC2</i>	alpha-(1,2)-fucosyltransferase Lewis antigen (VF0057)
	<i>gtrA</i>	bactoprenol-linked glucose translocase/flippase LPS (VF0124)
	<i>HL_RS04620</i>	lipopolysaccharide biosynthesis protein LOS (VF0044)
	<i>htrB</i>	lipid A biosynthesis lauroyl acyltransferase LPS (VF0370)
	<i>KPHS_35660</i>	glycosyl transferase family protein Capsule (VF0560)
	<i>KPN_RS13475</i>	acyltransferase Capsule (VF0560)
	<i>lsgE</i>	glycosyltransferase LOS (VF0044)
	<i>mce3D</i>	MCE family protein Mce3 (VF0826)
		virulence factor Mce family protein Mce3 (VF0826)
		Mce family protein Mce3D Mce3 (VF0826)
	<i>mce4D</i>	MCE-family protein Mce4 (VF0827)
		Mce family protein Mce4D Mce4 (VF0827)
	<i>PA3143</i>	hypothetical protein LPS (VF0085)
	<i>SGO_RS05030</i>	glycosyltransferase family 2 protein Capsule (VF0144)

	<i>SH_RS01850</i>	glycosyltransferase family 4 protein Capsule (VF0003)
	<i>SSA_RS10865</i>	glycosyltransferase Capsule (VF0144)
	<i>stf0</i>	Stf0 sulfotransferase family protein Sulfolipid-1 biosynthesis and transport (VF0803)
		hypothetical protein Sulfolipid-1 biosynthesis and transport (VF0803)
		Conserved hypothetical protein (sulfotransferase?) Sulfolipid-1 biosynthesis and transport (VF0803)
	<i>tesA</i>	alpha/beta fold hydrolase PDIM (VF0309)
Motility (VFC0204)	<i>YPTS_RS05430</i>	glycosyltransferase family 4 protein O-antigen (VF0392)
	<i>BCE_RS08880</i>	flagellar type III secretion system pore protein FlIP Peritrichous flagella (VF0670)
	<i>flaA</i>	flagellin Flagella (VF0114)
		flagellin A FlaA Flagella (VF0051)
	<i>flaB</i>	flagellin Flagella (VF0114)
		flagellin B FlaB Flagella (VF0051)
	<i>flgD</i>	flagellar hook capping protein Polar flagella (VF0473)
	<i>flgG</i>	flagellar basal-body rod protein FlgG Flagella (VF0051)
	<i>fliR</i>	flagellar biosynthetic protein FlIR Flagella (VF0430)
	<i>fliS</i>	flagellar export chaperone FlIS Flagella (VF0685)
Nutritional/Metabolic factor (VFC0272)	<i>motY</i>	sodium-type flagellar protein MotY Flagella (VF0519)
	<i>dhbB</i>	isochorismatase, Dhbb Bacillibactin (VF0586)
	<i>dhbC</i>	isochorismate synthase MenF Brucebactin (VF0692)
	<i>hemO</i>	biliverdin-producing heme oxygenase HemO cluster (VF0636)
	<i>iucB</i>	aerobactin siderophore biosynthesis protein IucB Aerobactin (VF0229)
		N-acetyltransferase IucB Aerobactin (VF0565)
		aerobactin synthesis protein IucB Aerobactin (VF0123)
	<i>mbtM</i>	Probable fatty acyl-AMP ligase MbtM Mycobactin (VF0299)
	<i>pdtorfE</i>	putative membrane protein Pyridine-2,6-dithiocarboxylic acid (PDTC) (VF0938)
	<i>phzM</i>	phenazine-specific methyltransferase PhzM (adenosylmethionine dependent methyltransferase) Pyocyanin (VF0100)
	<i>pvdG</i>	pyoverdine biosynthesis protein PvdG Pyoverdine (VF0094)
	<i>viuA</i>	TonB-dependent siderophore vulnibactin receptor VvuA Vibriobactin receptor ViuA (VF0640)
	<i>ybtU</i>	yersiniabactin biosynthesis oxidoreductase YbtU Ybt (VF0564)
		yersiniabactin biosynthetic protein YbtU Yersiniabactin (VF0136)
		yersiniabactin biosynthesis oxidoreductase YbtU Yersiniabactin siderophore (VF1115)
Regulation (VFC0301)	<i>acpA</i>	capsule synthesis transcriptional regulator AcpA AcpAB (VF0666)
Stress survival (VFC0282)	<i>mntC</i>	periplasmic binding protein MntC MntABC (VF0455)
Others (VFC0346)	<i>virK</i>	virulence factor VirK VirK (VF0980)
		hypothetical protein VirK (VF0980)

Annex 4. Virulence factors differentially abundant in CM compared to LB.

Category	VF gene name	Description	Base Mear	log2FC	p-adj
	<i>srtC-2/srtC</i>	sortase RlrA islet (VF0529)	1.973	2.002	0.002
	<i>PM_RS00425</i>	GspE/PuIE family protein Type IV pili (VF1229)	4.084	2.008	0.023
	<i>EFME1162_RS23495</i>	YSIRK signal domain/LPXTG anchor domain surface protein Esp (VF0353)	5.273	2.010	0.037
	<i>pilB</i>	type IV-A pilus assembly ATPase PilB TFP (VF1334)	2.235	2.143	0.003
			2.134	2.022	0.004
			2.169	2.022	0.004
			2.285	2.150	0.003
			2.091	2.042	0.003
		GspE/PuIE family protein Type IV pili (VF0753)	4.032	2.237	0.011
			4.032	2.237	0.011
		type IV-A pilus assembly ATPase PilB Type IV pili (VF1240)	4.580	2.034	0.027
	<i>pilH</i>	response regulator TFP (VF1334)	2.937	2.852	0.000
	<i>DIP_RS14950</i>	UPF0182 family protein DIP0733 (VF0601)	6.700	2.069	0.035
	<i>fbpA</i>	fibronectin-binding protein FbpA (VF0349)	7.008	2.081	0.045
			3.454	2.055	0.018
			3.982	2.056	0.026
			3.897	2.018	0.027
			3.625	2.140	0.015
			4.777	2.141	0.028
			3.756	2.160	0.014
			4.983	2.204	0.022
			4.276	2.213	0.018
			4.242	2.177	0.020
			5.072	2.218	0.022

Adherence (VFC0001)

		4.403	2.148	0.023
<i>bfpF</i>	retraction ATPase BFP (VF0174)	1.880	2.170	0.002
<i>peb1</i>	transporter substrate-binding domain-containing protein PEB1 (VF0772)	2.786	2.176	0.006
<i>sdrF</i>	YSIRK-type signal peptide-containing protein SDr (VF0019)	3.191	2.148	0.012
<i>gspB</i>	platelet binding protein GspB Serine-rich surface glycoproteins (VF1050)	1.888	2.177	0.002
<i>hcpB</i>	type II secretion system protein GspE Hemorrhagic E.coli pilus (HCP) (VF1141)	3.344	2.274	0.003
		2.791	2.038	0.007
		3.700	2.309	0.005
		3.518	2.351	0.003
		3.518	2.367	0.003
		5.410	3.215	0.000
		5.517	3.095	0.000
<i>fbpA/fbp68</i>	fibronectin-binding protein FbpA FbpA/Fbp68 (VF0595)	4.947	2.923	0.000
		5.285	2.489	0.010
		6.814	4.367	0.000
		4.993	4.474	0.000
<i>pilA</i>	pilin ChiRP (VF0609)	4.231	2.291	0.011
<i>pebA</i>	bifunctional adhesin/ABC transporter aspartate/glutamate-binding protein PEB1 (VF0327)	1.972	2.302	0.000
<i>xcpA/pilD</i>	type 4 prepilin peptidase PilD Type IV pili (VF0082)	4.260	2.314	0.009
<i>CPE_RS11865</i>	A24 family peptidase Type IV pili (VF0731)	2.844	2.438	0.002
<i>pilD</i>	type IV pilus prepilin peptidase Type IV pili (VF0075)	2.722	2.266	0.005
		2.801	2.373	0.003
		2.373	2.380	0.001
		2.373	2.380	0.001
		2.373	2.380	0.001
		2.258	2.431	0.001
		2.373	2.523	0.001
		3.929	2.831	0.001
		2.387	2.402	0.001
	type IV prepilin leader peptide type M1 Type IV pili (VF0431)	2.243	2.407	0.001
<i>clfB</i>	Clumping factor B; adhesin Clumping factor (VF0004)	4.161	2.702	0.002
<i>cshB</i>	LPXTG-anchored fibrillar adhesin CshB Cell surface hydrophobicity proteins (VF1049)	2.518	2.391	0.001
<i>CPE_RS09470</i>	GspE/PuE family protein Type IV pili (VF0731)	3.436	2.401	0.006
<i>cshA</i>	CshA/CshB family fibrillar adhesin-related protein Cell surface hydrophobicity proteins (VF1049)	4.694	2.415	0.010
		3.436	2.417	0.006
		3.450	2.368	0.007
	LPXTG-anchored fibrillar adhesin CshA Cell surface hydrophobicity proteins (VF1049)	3.608	2.368	0.007
<i>ami</i>	autolysin amidase; adhesin Ami (VF0071)	3.579	2.401	0.006
		2.958	2.317	0.006
		3.072	2.390	0.004
		3.115	2.458	0.003
		3.115	2.458	0.003
		3.115	2.458	0.003
		3.143	2.458	0.003
		3.143	2.458	0.003
		3.143	2.458	0.003
		3.071	2.407	0.004
		5.016	3.921	0.000
		3.100	2.441	0.003
		3.100	2.424	0.004
		3.086	2.424	0.004
<i>pilD/vcpD</i>	A24 family peptidase ChiRP (VF0609)	2.829	2.278	0.006
		2.630	2.113	0.009
		2.687	2.481	0.001
<i>fbp54</i>	fibronectin-bing protein Fbp54 FBPs (VF0243)	2.914	2.516	0.002
		7.374	2.348	0.026
		5.598	2.380	0.016
		7.653	2.273	0.033
		7.567	2.323	0.029
		7.208	2.414	0.022
		6.836	2.540	0.013
		5.337	2.542	0.009
		7.143	2.452	0.020
		7.071	2.492	0.018
		5.879	2.493	0.011
		7.007	2.495	0.017
		6.037	2.466	0.013
		7.063	2.472	0.018
		14.989	2.430	0.039
		7.362	2.487	0.018
		4.995	2.776	0.003

<i>tapD/pilD</i>	prepilin peptidase Tap type IV pili (VF0475)	2.978	2.516	0.002
		2.914	2.497	0.002
		2.928	2.497	0.002
<i>tcpJ</i>	prepilin signal peptidase TcpJ TCP (VF0126)	4.920	2.500	0.009
<i>cbpA/pspC</i>	choline binding protein A CbpA/PspC (VF0150)	3.177	2.912	0.000
		2.804	2.526	0.001
<i>pavA</i>	Fibronectin-binding protein-like protein A FBPs (VF0243)	8.322	2.539	0.018
		8.402	2.506	0.019
		8.104	2.553	0.017
		7.297	2.376	0.024
		6.656	2.134	0.039
		6.670	2.146	0.038
	Fibronectin-binding protein-like protein A PavA (VF0283)	12.784	2.300	0.042
<i>chpA</i>	still frameshift probable component of chemotactic signal transduction system Type IV pili (VF0082)	7.261	2.871	0.001
<i>MCR_RS01565</i>	type IV pilus twitching motility protein PilT Type IV pili (VF1228)	2.741	2.883	0.000
<i>lap</i>	Listeria adhesion protein Lap Lap (VF0444)	35.170	2.939	0.013
		35.170	2.939	0.013
		35.170	2.939	0.013
		35.170	2.939	0.013
		35.170	2.939	0.013
		35.170	2.939	0.013
		35.170	2.939	0.013
		153.071	5.216	0.000
<i>SGO_RS00800</i>	C69 family dipeptidase AbpB SGO_0162 (VF1070)	7.000	3.183	0.001
<i>hsa</i>	accessory Sec-dependent serine-rich glycoprotein adhesin SrpA SSA_0829 (VF1101)	2.057	2.673	0.000
	accessory Sec-dependent serine-rich glycoprotein adhesin Serine-rich surface glycoproteins (VF1050)	3.212	3.127	0.000
<i>sdrE</i>	Ser-Asp rich fibrinogen-binding bone sialoprotein-binding protein SDR (VF0019)	5.391	3.135	0.000
		6.496	2.571	0.011
		3.308	2.167	0.012
		3.351	2.167	0.012
		3.600	2.247	0.010
		3.358	2.179	0.012
		3.708	2.328	0.007
		3.365	2.215	0.010
		3.280	2.144	0.013
<i>AAM90673</i>	surface protein SdrI SDR (VF0019)	4.400	3.839	0.000
<i>CD0873</i>	ABC transporter substrate-binding protein CD0873 (VF0593)	5.833	2.952	0.002
<i>DIP_RS19245</i>	C40 family peptidase DIP1621 (VF0603)	6.839	3.374	0.000
<i>aaa</i>	autolysin/adhesin Aaa SACOL0507 (VF0999)	4.501	4.195	0.000
<i>aae</i>	autolysin/adhesin Aae SE2319 (VF1000)	5.046	4.350	0.000
<i>pilT</i>	twitching motility protein PilT Type IV pili (VF0082)	1.856	2.578	0.000
		3.128	3.533	0.000
<i>atlC</i>	fibronectin binding autolysin Autolysin alt (VF0997)	7.672	4.468	0.000
<i>sdrC</i>	Ser-Asp rich fibrinogen-binding bone sialoprotein-binding protein SDR (VF0019)	4.450	2.686	0.001
		9.198	4.484	0.000
		9.363	4.504	0.000
		9.406	4.505	0.000
		9.406	4.505	0.000
		8.163	4.598	0.000
		10.231	4.651	0.000
		8.270	4.341	0.000
<i>sdrD</i>	Ser-Asp rich fibrinogen-binding bone sialoprotein-binding protein SDR (VF0019)	21.500	5.021	0.000
		6.306	4.203	0.000
		6.256	4.203	0.000
		7.082	4.447	0.000
		7.765	4.570	0.000
		7.345	4.513	0.000
<i>atl</i>	bifunctional autolysin Autolysin alt (VF0997)	7.609	4.541	0.000
		13.989	4.870	0.000
		4.324	4.268	0.000
		6.462	4.592	0.000
		5.257	4.552	0.000
		7.087	4.651	0.000
		7.116	4.655	0.000
		6.105	4.685	0.000
	glucosaminidase domain-containing protein Autolysin alt (VF0997)	8.742	4.834	0.000
		8.742	4.834	0.000
		8.763	4.839	0.000
		8.763	4.839	0.000
	GW dipeptide domain-containing protein Autolysin alt (VF0997)	4.750	4.423	0.000
		4.059	3.088	0.000
		9.307	4.851	0.000

Antimicrobial activity/Competitive advantage (VFC0325)	<i>farB</i>	fatty acid efflux system protein FarB FarAB (VF0450)	6.237	2.289	0.018
			6.252	2.289	0.018
			10.475	2.398	0.020
Biofilm (VFC0271)	<i>icaA</i>	N-acetylglucosaminyltransferase; involved in polysaccharide intercellular adhesin(PIA Intercellular adhesion proteins (VF0014)	1.811	2.019	0.001
			5.577	2.089	0.031
	<i>bopD</i>	sugar-binding transcriptional regulator; LacI family BopD (VF0362)	5.031	2.144	0.023
			5.031	2.144	0.023
			10.675	2.189	0.048
			1.928	2.506	0.000
	<i>fsrA</i>	response regulator Fsr (VF0360)	2.916	3.412	0.000
			1.928	2.506	0.000
	<i>vpsC</i>	exopolysaccharide biosynthesis acetyltransferase VpsC VPS (VF0615)	7.384	2.834	0.005
			10.186	3.666	0.000
	<i>mucD</i>	serine protease MucD precursor Alginate (VF0091)	10.075	3.721	0.000
			7.897	3.728	0.000
			10.653	3.765	0.000
			10.697	3.766	0.000
			10.671	3.789	0.000
	<i>pgaC</i>	poly-beta-1,6 N-acetyl-D-glucosamine synthase PNAG (VF0472)	3.931	3.831	0.000
			3.917	3.831	0.000
			3.889	3.865	0.000
			7.059	4.294	0.000
	<i>algW</i>	AlgW protein Alginate (VF0091)	10.085	4.168	0.000
			10.335	4.177	0.000
			10.157	4.177	0.000
			10.624	4.182	0.000
			20.216	4.260	0.000
			8.406	3.476	0.000
Effector delivery system (VFC0086)	<i>aec27/clpV</i>	type VI secretion system ATPase TssH ACE T6SS (VF1122)	5.203	2.012	0.031
			7.268	2.095	0.033
			8.212	2.078	0.038
			5.792	2.140	0.021
			7.540	2.056	0.038
			9.631	2.109	0.040
	<i>ssaN</i>	type III secretion system ATPase SsaN TTSS (SPI-2 encode) (VF0321)	3.099	2.014	0.014
			3.142	2.014	0.014
			5.746	2.142	0.026
	<i>gspE2</i>	putative secretion pathway ATPase T2SS (VF1335)	5.457	2.579	0.004
			12.421	4.520	0.000
			5.570	2.370	0.011
			5.670	2.413	0.010
			5.671	2.458	0.008
			23.758	4.750	0.000
	<i>lspE</i>	general secretion pathway protein E Lsp T2SS (VF0154)	3.937	2.674	0.001
			4.265	2.853	0.000
			8.260	3.055	0.001
			4.127	2.827	0.000
	<i>clpV/tssH</i>	type VI secretion system ATPase TssH T6SS (VF1337)	4.503	2.330	0.007
			7.009	2.158	0.029
			7.626	2.176	0.030
			7.323	2.187	0.027
		type VI secretion system ATPase TssH T6SS (VF0569)	7.271	2.250	0.023
	<i>sctN</i>	type VI secretion system ATPase TssH T6SS (VF1265)	15.254	2.273	0.039
		type III secretion system ATPase SctN T3SS (VF1265)	4.603	2.144	0.017
		type III secretion system ATPase SctN T3SS (VF1274)	3.821	2.151	0.014
		type III secretion system ATPase SctN T3SS (VF1270)	6.906	3.825	0.000
	<i>tssH</i>	type VI secretion system ATPase TssH HSI-3 (VF0944)	9.884	2.066	0.046
		type VI secretion system ATPase TssH HSI-3 (VF0944)	21.981	3.991	0.000
		type VI secretion system ATPase TssH SCI-I T6SS (VF1120)	33.428	4.030	0.000
		type VI secretion system ATPase TssH CTS2 (VF1264)	33.449	4.651	0.000
		type VI secretion system ATPase TssH T6SS-A (VF1036)	6.496	2.051	0.033
		type VI secretion system ATPase TssH Imp (VF1258)	6.698	2.157	0.029
		type VI secretion system ATPase TssH HSI-2 (VF0943)	11.748	2.196	0.039
		type VI secretion system ATPase TssH CTS1 (VF1263)	14.126	2.364	0.032
	<i>xcpR</i>	general secretion pathway protein E Xcp secretion system (VF0084)	8.559	2.081	0.042
			2.842	2.167	0.006
	<i>eccA3</i>	Type VII secretion system protein EccA3 ESX-3 (VF0521)	2.899	2.098	0.009
			2.871	2.135	0.008
	<i>legK1</i>	Dot/Icm type IV secretion system effector LegK1; serine/threonine-protein kinase Dot/Icm T4SS secreted effectors (VF0798)	2.823	2.127	0.007
	<i>gspE1</i>	general secretion pathway protein E T2SS (VF1335)	4.951	2.014	0.026
	<i>SG_RS05240</i>	AAA family ATPase T6SS (VF0975)	5.728	2.323	0.013
	<i>glgX</i>	glycogen hydrolase TTSS secreted effectors (VF0711)	21.513	3.594	0.000

	<i>NGR_RS11050</i>	Flii/YscN family ATPase T3SS2 (VF1299)	4.884	2.349	0.007
	<i>CPN_RS04405</i>	Flii/YscN family ATPase TTSS (VF0344)	2.119	2.362	0.001
	<i>clpB</i>	type VI secretion system ATPase ClpV1 T6SS (VF0480)	9.084	2.488	0.013
			20.377	3.211	0.003
	<i>CT_473</i>	hypothetical protein TTSS secreted effectors (VF0711)	3.783	2.636	0.001
	<i>yst10</i>	A24 family peptidase Yst1 T2SS (VF0635)	4.630	2.846	0.002
	<i>fill</i>	type III secretion system ATPase TTSS (VF0344)	2.974	2.894	0.000
	<i>eccCa1</i>	ESX-1 type VII secretion system FtsK/SpoIIIE family protein EccCa1 ESX-1 (VF0315)	3.059	3.023	0.000
			2.313	2.862	0.000
	<i>pkn5</i>	serine/threonine protein kinase TTSS (VF0344)	12.676	3.168	0.001
			12.676	3.168	0.001
	<i>cdsN</i>	Type III secretion system ATPase TTSS (VF0344)	6.200	3.352	0.000
			6.157	3.377	0.000
			8.103	3.841	0.000
			6.279	3.496	0.000
			11.671	3.568	0.000
			7.526	3.665	0.000
			7.468	3.665	0.000
			7.554	3.665	0.000
	<i>hrcN</i>	type III secretion system ATPase HrcN P, syringae TTSS (VF0909)	7.338	3.848	0.000
			7.353	3.853	0.000
		HrcN T3SS (VF1283)	9.080	4.290	0.000
		Flii/YscN family ATPase P, syringae TTSS (VF0909)	7.424	3.877	0.000
	<i>essC</i>	type VII secretion system protein EssC T7SS (VF0581)	14.190	3.461	0.000
			6.583	3.242	0.000
			7.296	3.349	0.000
		type VII secretion system protein EssC; FtsK/SpoIIIE family ATPase Type VII secretion system (VF0403)	6.226	3.873	0.000
			6.205	3.907	0.000
			6.205	3.907	0.000
			6.233	3.845	0.000
			11.765	4.141	0.000
	<i>tagT</i>	type VI secretion associated protein TagT; ATP-binding component of ABC transporter HSI-1 (VF0334)	40.356	3.923	0.001
	<i>clpV1</i>	type VI secretion system AAA+ family ATPase HSI-1 (VF0334)	26.962	4.087	0.000
	<i>ETAE_RS04185</i>	Flii/YscN family ATPase T3SS (VF1266)	7.686	4.263	0.000
	<i>tssH/clpV1</i>	type VI secretion system ATPase TssH Putida K1-T6SS (VF1351)	25.858	4.292	0.000
			21.788	4.326	0.000
			28.464	4.412	0.000
			26.246	4.269	0.000
Exoenzyme (VFC0251)	<i>geh</i>	glycerol ester hydrolase Lipase (VF0012)	1.618	2.016	0.000
	<i>nagJ</i>	hyaluronidase Mu-toxin (VF0389)	2.128	2.070	0.001
	<i>aur</i>	zinc metalloproteinase aureolysin Aureolysin (VF0024)	2.906	2.117	0.009
	<i>mf3</i>	deoxyribonuclease Mitogenic factor (VF0252)	3.579	2.252	0.008
	<i>htrA/degP</i>	S1C family serine protease Serine protease (VF1055)	9.075	2.536	0.013
			9.075	2.536	0.013
			9.583	2.398	0.026
		trypsin-like peptidase domain-containing protein Serine protease (VF1055)	7.734	2.456	0.014
			9.136	2.456	0.020
			9.144	2.456	0.020
			9.144	2.456	0.020
			8.866	2.491	0.014
			9.090	2.543	0.013
			9.100	2.560	0.011
			8.128	2.574	0.010
	<i>mf2</i>	deoxyribonuclease; phage associated Mitogenic factor (VF0252)	1.944	2.672	0.000
	<i>eno</i>	phosphopyruvate hydratase Streptococcal enolase (VF1060)	37.787	3.475	0.002
			37.778	3.476	0.002
			37.821	3.477	0.002
	<i>zmp1</i>	M13 family metallopeptidase Zn++ metalloprotease (VF0846)	37.821	3.477	0.002
			18.359	4.321	0.000
			17.711	4.485	0.000
			16.891	4.545	0.000
			16.891	4.545	0.000
			16.798	4.588	0.000
			25.371	4.597	0.000
			16.958	4.619	0.000
			16.931	4.655	0.000
			23.320	4.687	0.000
			23.378	4.692	0.000
			23.378	4.692	0.000
			23.378	4.692	0.000
			25.371	4.597	0.000
			23.372	4.705	0.000
			22.471	4.739	0.000
			24.767	4.798	0.000
			22.494	4.826	0.000

			22.508	4.826	0.000	
		Probable zinc metalloprotease Zn++ metalloprotease (VF0846)	16.929	4.618	0.000	
		zinc metalloprotease Zn++ metalloprotease (VF0846)	21.001	4.734	0.000	
			16.677	4.604	0.000	
			23.589	4.803	0.000	
			23.589	4.803	0.000	
		zinc metalloprotease Zmp1 Zn++ metalloprotease (VF0846)	23.589	4.803	0.000	
		23.589	4.803	0.000		
Exotoxin (VFC0235)	cylZ	3R-hydroxymyristoyl ACP dehydratase CylZ -haemolysin/cytolysin (VF0279)	2.605	-2,002	0.034	
	cylG	3-ketoacyl-ACP-reductase CylG -haemolysin/cytolysin (VF0279)	7.765	2.097	0.039	
	ASA_RS04230	hemolysin III Hemolysin III (VF0647)	3.364	2.172	0.011	
	AHA_RS17655	hemolysin III Hemolysin III (VF0647)	3.320	2.211	0.009	
	AHML_RS18555	hemolysin III Hemolysin III (VF0647)	3.334	2.211	0.009	
	B565_RS04085	hemolysin III Hemolysin III (VF0647)	3.291	2.230	0.008	
	hlyIII	putative membrane hydrolase Hemolysin III (VF0655)	3.445	2.236	0.009	
		hemolysin III family protein Hemolysin III (VF0655)	3.830	2.316	0.009	
	rtxE	RTX toxin transporter; ATPase protein RtxA (VF0482)	35.874	2.470	0.044	
	cyaB	cyclolysin secretion ATP-binding protein Cya (VF0028)	18.188	2.391	0.036	
	hlyA	hemolysin A Hemolysin (VF0879)	4.413	2.418	0.007	
		TlyA family RNA methyltransferase Hemolysin (VF0879)	4.441	2.418	0.007	
			2.302	2.357	0.001	
			4.189	2.477	0.004	
			3.765	2.546	0.002	
	clbH	colibactin non-ribosomal peptide synthetase ClbH Colibactin (VF1179)	3.915	2.600	0.002	
		colibactin non-ribosomal peptide synthetase ClbH Colibactin (VF0573)	7.046	2.944	0.002	
	rtxB	RTX toxin transporter; ATPase protein RtxA (VF0482)	21.026	3.225	0.003	
	cmaA	L-allo-isoleucine:holo-[CmaA peptidyl-carrier protein] ligase Phytotoxin coronatine (VF0916)	5.094	3.753	0.000	
		galF	GalU regulator GalF Capsule (VF0560)	4.752	2.005	0.023
				4.665	2.005	0.023
4.781				2.005	0.023	
8.556				2.169	0.030	
JJD26997_RS08105		gamma-glutamyl-gamma-aminobutyrate hydrolase family protein Capsule (VF0323)	1.986	2.008	0.003	
KPHS_35740		UDP-glucose pyrophosphorylase Capsule (VF0560)	4.187	2.012	0.020	
FTN_RS06220		glycosyltransferase family 2 protein Capsule (VF0543)	5.609	2.023	0.037	
FTW_RS02930		glycosyltransferase family 2 protein Capsule (VF0543)	5.560	2.029	0.035	
FTA_RS07270		glycosyltransferase family 2 protein Capsule (VF0543)	5.890	2.033	0.037	
STER_RS05280		glycosyltransferase family 1 protein Capsule (VF0144)	1.614	2.036	0.001	
YE105_RS06200		glycosyltransferase family 2 protein O-antigen (VF0392)	1.886	2.048	0.001	
KOX_RS24940		DUF4422 domain-containing protein LPS (VF0561)	5.801	2.050	0.041	
BT9727_RS25850		sugar transferase Polysaccharide capsule (VF0659)	6.369	2.062	0.034	
pbpG		D-alanyl-D-alanine endopeptidase PBP7/8 PbpG (VF0468)	3.873	2.065	0.021	
			3.873	2.065	0.021	
			3.873	2.065	0.021	
			3.873	2.065	0.021	
			5.982	2.278	0.018	
		D-alanyl-D-alanine carboxypeptidase PbpG (VF0468)	3.873	2.065	0.021	
D-alanyl-D-alanine endopeptidase PbpG (VF0468)		3.873	2.065	0.021		
pknG		serine/threonine-protein kinase PknG Protein kinase G (VF0805)	2.855	2.077	0.012	
			2.855	2.077	0.012	
			2.812	2.078	0.011	
			2.797	2.078	0.011	
			2.871	2.102	0.012	
			2.769	2.079	0.010	
			2.796	2.013	0.015	
			2.796	2.013	0.015	
			2.943	2.125	0.011	
			3.016	2.172	0.009	
			2.927	2.035	0.013	
			2.739	2.042	0.014	
		serine/threonine protein kinase Protein kinase G (VF0805)	22.551	3.424	0.001	
lpxA/glmU		UDP-N-acetylglucosamine pyrophosphorylase/glucosamine-1-phosphate N-acetyltransferase LPS (VF0542)	5.029	2.236	0.013	
			8.806	2.237	0.033	
			4.987	2.246	0.013	
			4.872	2.190	0.014	
			8.850	2.085	0.048	
			4.915	2.350	0.008	
STU_RS14630		sugar transferase Capsule (VF0144)	9.445	2.355	0.022	
			5.731	2.087	0.034	
			8.331	2.092	0.037	
	7.741		2.108	0.033		
		6.811	2.109	0.029		

oatA	peptidoglycan O-acetyltransferase OatA (VF0441)	6.811	2.109	0.029
		6.889	2.122	0.028
		6.889	2.122	0.028
		6.889	2.122	0.028
		7.893	2.063	0.037
		6.709	2.074	0.033
		7.336	2.153	0.027
		7.336	2.153	0.027
		7.358	2.153	0.027
		13.216	2.238	0.038
SPD_RS01735	sugar transferase Capsule (VF0144)	5.209	2.111	0.022
manA	mannose-6-phosphate isomerase; class I Polysaccharide capsule (VF0659)	6.540	2.113	0.035
SPH_RS02230	sugar transferase Capsule (VF0144)	5.266	2.121	0.026
BT9727_RS25860	tyrosine protein phosphatase Polysaccharide capsule (VF0659)	4.976	2.128	0.029
FTM_RS05910	glycosyltransferase family 2 protein Capsule (VF0543)	5.977	2.129	0.027
galU	glucosephosphate uridylyltransferase LOS (VF0044)	6.188	2.234	0.024
		6.271	2.150	0.031
acpXL	acyl carrier protein LPS (VF0367)	5.262	2.145	0.025
C8J_RS08455	glycosyltransferase Capsule (VF0323)	6.398	2.151	0.032
flmF2	glycosyl transferase family protein LPS (VF0542)	3.772	2.140	0.016
		3.763	2.140	0.016
		3.915	2.228	0.012
		4.028	2.299	0.010
		6.245	3.243	0.000
SPCG_RS01825	sugar transferase Capsule (VF0144)	5.167	2.175	0.020
BC_RS26260	glycosyltransferase Polysaccharide capsule (VF0659)	3.738	2.175	0.012
wbaP	undecaprenyl-phosphate galactose phosphotransferase WbaP Capsule (VF0560)	5.992	2.176	0.025
		5.351	2.018	0.037
		6.103	2.249	0.019
		6.104	2.269	0.018
		6.588	2.338	0.017
FTA_RS07275	glycosyltransferase family 2 protein Capsule (VF0543)	7.475	2.210	0.028
STER_RS05300	sugar transferase Capsule (VF0144)	5.739	2.212	0.024
waaV	glucosyltransferase LOS (VF0326)	4.246	2.214	0.020
SAK_RS06290	glycosyltransferase Capsule (VF0274)	3.986	2.218	0.011
KPHS_35670	UDP-Gal::undecaprenolphosphate Gal-1-P transferase Capsule (VF0560)	6.074	2.223	0.022
SMU_RS01255	rhamnose-glucose polysaccharide biosynthesis protein RgpB Capsule (VF0144)	8.714	2.225	0.029
wzt2	ATP-binding ABC transporter capsular polysaccharide export protein Capsule I (VF0436)	9.296	2.600	0.013
		7.906	2.232	0.024
bexA	ATP-dependent polysaccharide export protein BexA Capsule (VF0043)	15.470	2.233	0.047
FTN_RS06225	glycosyltransferase family 2 protein Capsule (VF0543)	7.562	2.236	0.026
FN3523_RS06380	glycosyltransferase family 2 protein Capsule (VF0543)	6.056	2.166	0.026
bpsD	bacterial sugar transferase family protein B, cereus exo-polysaccharide (BPS) (VF0668)	5.621	2.246	0.019
wbfU	sugar transferase Capsular polysaccharide (VF0624)	2.128	2.415	0.000
		2.272	2.238	0.001
wlaN	beta-1,3 galactosyltransferase LOS (VF0326)	4.136	2.271	0.018
C8J_RS05610	glycosyltransferase family 2 protein LOS (VF0326)	3.182	2.275	0.005
BALH_RS26235	sugar transferase Polysaccharide capsule (VF0659)	5.910	2.279	0.017
SSU98_0572	Cps2E Capsule (VF0144)	5.203	2.283	0.014
F92_RS07260	glycosyltransferase family 2 protein Capsule (VF0543)	6.961	2.296	0.019
SSU05_0568	Cps2E Capsule (VF0144)	3.609	2.307	0.009
STR_RS05225	acyltransferase Capsule (VF0144)	2.886	2.313	0.005
bpsX	LytR family transcriptional regulator B, cereus exo-polysaccharide (BPS) (VF0668)	9.177	2.318	0.033
CJJ81176_RS06815	gamma-glutamyl-gamma-aminobutyrate hydrolase family protein Capsule (VF0323)	2.572	2.334	0.002
wzc	polysaccharide biosynthesis tyrosine autokinase Capsular polysaccharide (VF0624)	3.918	2.339	0.012
rib	alpha-like surface protein Rib (VF1064)	2.147	2.358	0.001
bpsB	capsular exopolysaccharide family domain protein B, cereus exo-polysaccharide (BPS) (VF0668)	3.697	2.368	0.011
STR_RS05250	tyrosine-protein kinase Capsule (VF0144)	3.279	2.371	0.006
C8J_RS08560	glycosyltransferase family 8 protein Capsule (VF0323)	3.368	2.392	0.007
mynA/sacA	UDP-N-acetylglucosamine 2-epimerase (non-hydrolyzing) Capsule (VF0079)	12.070	2.394	0.036
SSA_RS10885	tyrosine-protein kinase Capsule (VF0144)	3.176	2.394	0.004
SAG_RS08580	sugar transferase Capsule (VF0274)	6.563	2.441	0.012
SAK_RS06310	sugar transferase Capsule (VF0274)	6.563	2.441	0.012
rfbD	GDP-D-mannose dehydratase LPS (VF0056)	7.996	2.449	0.016
KPHS_35590	6-phosphogluconate dehydrogenase Capsule (VF0560)	22.676	2.457	0.033
		45.036	2.494	0.040

Immune modulation
(VFC0258)

<i>gndA</i>	NADP-dependent phosphogluconate dehydrogenase Capsule (VF0560)	22.068	2.511	0.028
		21.988	2.448	0.033
		22.691	2.460	0.033
		22.691	2.460	0.033
		22.691	2.460	0.033
		22.691	2.460	0.033
		22.691	2.460	0.033
		22.691	2.460	0.033
<i>orfM</i>	deoxyribonucleotide triphosphate pyrophosphatase LOS (VF0044)	6.733	2.465	0.009
		6.733	2.465	0.009
		6.733	2.465	0.009
		12.838	2.590	0.015
<i>adsA</i>	Adenosine synthase A AdsA (VF0422)	3.277	2.488	0.001
		3.277	2.488	0.001
		3.277	2.488	0.001
		5.707	2.698	0.002
		3.114	2.329	0.002
<i>gtrB</i>	bactoprenol glucosyl transferase LPS (VF0124)	9.341	3.217	0.001
		5.790	2.486	0.012
<i>SAOUHSC_00129</i>	UDP-N-acetylglucosamine 2-epimerase Capsule (VF0003)	11.111	2.491	0.024
<i>cpsD</i>	polysaccharide biosynthesis tyrosine autokinase Capsular polysaccharide (VF0624)	5.425	2.527	0.009
		4.400	2.554	0.003
<i>SMU_RS03835</i>	ABC transporter ATP-binding protein Capsule (VF0144)	6.273	2.525	0.007
<i>BCE_RS28105</i>	glycosyltransferase family 2 protein Polysaccharide capsule (VF0659)	3.625	2.549	0.002
<i>STER_RS07055</i>	ABC transporter ATP-binding protein Capsule (VF0144)	6.809	2.549	0.008
<i>STR_RS06895</i>	ABC transporter ATP-binding protein Capsule (VF0144)	6.809	2.549	0.008
<i>SAK_RS06335</i>	LysR family transcriptional regulator Capsule (VF0274)	6.594	2.575	0.007
<i>F7308_RS05570</i>	glycosyltransferase family 2 protein Capsule (VF0543)	3.794	2.576	0.001
<i>FNFX1_RS06220</i>	glycosyltransferase family 2 protein Capsule (VF0543)	6.572	2.585	0.007
<i>msbA</i>	lipid transporter ATP-binding/permease LOS (VF0044)	25.956	2.578	0.024
		25.978	2.578	0.024
		49.690	2.677	0.028
		26.998	2.637	0.021
		26.604	2.650	0.021
		26.899	2.663	0.020
		56.750	4.582	0.000
<i>FPHI_RS07665</i>	glycosyltransferase family 2 protein Capsule (VF0543)	9.038	2.686	0.010
<i>cps4D</i>	capsular polysaccharide biosynthesis protein Cps4D Capsule (VF0144)	5.097	2.668	0.006
<i>ABBFA_RS17100</i>	phosphomannomutase/phosphoglucomutase Capsule (VF0465)	8.201	2.710	0.005
<i>SPCG_RS01845</i>	glycosyltransferase family 2 protein Capsule (VF0144)	2.043	2.723	0.000
<i>BCAH187_RS26500</i>	glycosyltransferase family 2 protein Polysaccharide capsule (VF0659)	7.334	2.726	0.005
<i>Cj1417c</i>	amidotransferase Capsule (VF0323)	4.336	2.743	0.002
<i>A1S_0066</i>	hypothetical protein Capsule (VF0465)	8.008	2.749	0.004
<i>AB57_RS00570</i>	phosphomannomutase/phosphoglucomutase Capsule (VF0465)	7.914	2.750	0.004
<i>F7308_RS05565</i>	glycosyltransferase family 2 protein Capsule (VF0543)	8.482	2.760	0.006
<i>M3Q_RS01530</i>	phosphomannomutase/phosphoglucomutase Capsule (VF0465)	8.271	2.762	0.004
<i>BJAB07104_RS00555</i>	phosphomannomutase/phosphoglucomutase Capsule (VF0465)	8.828	2.764	0.005
<i>ABSD_F_RS00410</i>	phosphomannomutase/phosphoglucomutase Capsule (VF0465)	8.767	2.764	0.005
<i>ABK1_RS00550</i>	phosphomannomutase/phosphoglucomutase Capsule (VF0465)	8.433	2.769	0.004
<i>BJAB0715_RS00595</i>	phosphomannomutase/phosphoglucomutase Capsule (VF0465)	7.549	2.786	0.003
<i>ABTJ_RS18490</i>	phosphomannomutase/phosphoglucomutase Capsule (VF0465)	8.267	2.789	0.004
<i>cap8C</i>	type 8 capsular polysaccharide synthesis protein Cap8C Capsule (VF0003)	4.655	2.796	0.002
<i>SAG_RS08585</i>	tyrosine-protein kinase Capsule (VF0274)	5.797	2.800	0.004
<i>CJJ81176_RS06900</i>	glycosyltransferase family 2 protein Capsule (VF0323)	2.007	2.806	0.000
<i>CJJ81176_RS06895</i>	glycosyltransferase family 2 protein Capsule (VF0323)	2.160	2.812	0.000
<i>SSU05_0574</i>	cell wall biosynthesis glycosyltransferase Capsule (VF0144)	6.225	2.824	0.002
<i>F7308_RS05555</i>	ABC transporter ATP-binding protein Capsule (VF0543)	3.588	2.828	0.000
<i>ACICU_RS00500</i>	phosphomannomutase/phosphoglucomutase Capsule (VF0465)	16.984	2.848	0.010
<i>SPCG_RS01850</i>	glycosyltransferase Capsule (VF0144)	6.216	2.854	0.002
<i>FPHI_RS07645</i>	ABC transporter ATP-binding protein Capsule (VF0543)	3.496	2.859	0.000
<i>rfbK2</i>	O9 family phosphomannomutase RfbK2 LPS (VF0561)	8.970	2.860	0.004
		8.969	2.864	0.004
<i>FTN_RS06240</i>	ABC transporter ATP-binding protein Capsule (VF0543)	3.451	2.874	0.000
<i>cpsG</i>	MurB family protein Capsule (VF0361)	5.491	2.873	0.001
	O9 family phosphomannomutase RfbK1 LPS (VF0561)	9.173	2.885	0.003
	glycosyltransferase family 2 protein Capsule (VF0144)	3.849	2.896	0.000
<i>STR_RS05220</i>	glycosyltransferase family 2 protein Capsule (VF0144)	3.849	2.896	0.000
<i>BC_RS26250</i>	LytR family transcriptional regulator Polysaccharide capsule (VF0659)	15.022	2.908	0.009
<i>A225_RS18990</i>	DUF4422 domain-containing protein LPS (VF0561)	5.255	2.908	0.002
<i>FNFX1_RS06240</i>	ABC transporter ATP-binding protein Capsule (VF0543)	3.394	2.931	0.000
<i>rfbK1</i>	O9 family phosphomannomutase RfbK1 Capsule (VF0560)	17.524	2.938	0.008
<i>CJE_RS06375</i>	glycosyltransferase family 2 protein LOS (VF0326)	7.060	2.957	0.003

<i>GBS_RS06570</i>	glycosyltransferase Capsule (VF0274)	7.786	2.969	0.002
<i>BT9727_RS25810</i>	LytR family transcriptional regulator Polysaccharide capsule (VF0659)	10.891	2.974	0.006
<i>BC_RS26305</i>	tyrosine protein phosphatase Polysaccharide capsule (VF0659)	4.631	2.978	0.001
<i>ctrD</i>	capsule polysaccharide export ATP-binding protein CtrD Capsule (VF0079)	9.649	2.998	0.002
		18.794	3.115	0.005
<i>KPN_RS13395</i>	ABC transporter ATP-binding protein LPS (VF0561)	9.649	2.998	0.002
<i>FTA_RS07295</i>	ABC transporter ATP-binding protein Capsule (VF0543)	12.442	3.003	0.003
<i>FTM_RS05935</i>	ABC transporter ATP-binding protein Capsule (VF0543)	3.454	3.003	0.000
<i>FTH_RS07330</i>	ABC transporter ATP-binding protein Capsule (VF0543)	3.424	3.010	0.000
<i>BCE_RS25805</i>	LytR family transcriptional regulator Polysaccharide capsule (VF0659)	3.365	3.013	0.000
<i>BCAH187_RS26490</i>	LytR family transcriptional regulator Polysaccharide capsule (VF0659)	13.753	3.043	0.005
<i>rfaJ</i>	alpha-1;6 Glc transferase LPS (VF0056)	13.778	3.073	0.005
<i>ddrA</i>	ATP-binding cassette domain-containing protein PDIM (VF0309)	8.493	3.103	0.003
<i>alp2</i>	alpha-like surface protein Alpha-like protein (VF1063)	26.878	3.197	0.005
<i>GBS_RS06590</i>	sugar transferase Capsule (VF0274)	2.943	3.335	0.000
<i>STER_RS07090</i>	glycosyltransferase Capsule (VF0144)	11.097	3.341	0.001
<i>SSA_RS07425</i>	glycosyltransferase Capsule (VF0144)	6.589	3.343	0.000
<i>rpe</i>	ribulose-phosphate 3-epimerase Capsule (VF0543)	4.353	3.415	0.000
		9.285	3.403	0.000
		9.292	3.405	0.000
		9.263	3.435	0.000
		9.222	3.445	0.000
		9.222	3.445	0.000
<i>wecB</i>	UDP-N-acetylglucosamine 2-epimerase (non-hydrolyzing) Capsule (VF0144)	17.830	3.564	0.000
	UDP-N-acetylglucosamine 2-epimerase (non-hydrolyzing) Polysaccharide capsule (VF0659)	14.902	2.626	0.022
	UDP-N-acetylglucosamine 2-epimerase (non-hydrolyzing) LPS (VF0085)	5.318	3.340	0.000
<i>SGO_RS09905</i>	tyrosine-protein kinase Capsule (VF0144)	12.966	3.466	0.001
<i>FTT_RS04120</i>	ABC transporter ATP-binding protein Capsule (VF0543)	2.767	3.470	0.000
<i>wbjD/wecB</i>	UDP-N-acetylglucosamine 2-epimerase (non-hydrolyzing) Capsular polysaccharide (VF0624)	6.211	3.474	0.000
	UDP-N-acetylglucosamine 2-epimerase (non-hydrolyzing) Capsular polysaccharide (VF0624)	13.444	3.478	0.001
<i>cps4I</i>	capsular polysaccharide biosynthesis protein Cps4I Capsule (VF0144)	13.444	3.478	0.001
<i>cpsC</i>	teichoic acid biosynthesis protein; putative Capsule (VF0361)	12.253	2.285	0.045
<i>BC_RS26320</i>	CpsD/CapB family tyrosine-protein kinase Polysaccharide capsule (VF0659)	27.888	3.564	0.002
<i>BC_RS26310</i>	CpsD/CapB family tyrosine-protein kinase Polysaccharide capsule (VF0659)	10.673	3.599	0.000
<i>BCE_RS25880</i>	CpsD/CapB family tyrosine-protein kinase Polysaccharide capsule (VF0659)	4.958	3.648	0.000
<i>PA3157</i>	acetyltransferase LPS (VF0085)	5.861	3.683	0.000
<i>Cj1136</i>	glycosyltransferase LOS (VF0326)	5.861	3.683	0.000
<i>BT9727_RS25865</i>	CpsD/CapB family tyrosine-protein kinase Polysaccharide capsule (VF0659)	11.510	3.700	0.000
<i>BALH_RS26250</i>	CpsD/CapB family tyrosine-protein kinase Polysaccharide capsule (VF0659)	4.553	3.732	0.000
<i>BCAH187_RS26565</i>	CpsD/CapB family tyrosine-protein kinase Polysaccharide capsule (VF0659)	5.832	3.749	0.000
<i>BCAH187_RS26575</i>	CpsD/CapB family tyrosine-protein kinase Polysaccharide capsule (VF0659)	5.832	3.749	0.000
<i>BCE_RS25890</i>	CpsD/CapB family tyrosine-protein kinase Polysaccharide capsule (VF0659)	5.989	3.749	0.000
<i>cap8B</i>	type 8 capsular polysaccharide synthesis protein Cap8B Capsule (VF0003)	5.215	3.756	0.000
		5.343	3.768	0.000
		8.941	4.229	0.000
		4.890	3.774	0.000
		4.890	3.774	0.000
<i>cap8P</i>	type 8 capsular polysaccharide synthesis protein Cap8P Capsule (VF0003)	4.890	3.774	0.000
		20.045	3.473	0.002
		11.097	2.491	0.024
		11.111	2.491	0.024
		11.111	2.491	0.024
<i>BT9727_RS25875</i>	CpsD/CapB family tyrosine-protein kinase Polysaccharide capsule (VF0659)	11.111	2.491	0.024
<i>BALH_RS26260</i>	CpsD/CapB family tyrosine-protein kinase Polysaccharide capsule (VF0659)	5.329	3.775	0.000
<i>KOX_RS24945</i>	ABC transporter ATP-binding protein LPS (VF0561)	5.329	3.775	0.000

	<i>A225_RS18995</i>	ABC transporter ATP-binding protein LPS (VF0561)	19.492	3.833	0.000
	<i>KPHS_35490</i>	lipopolysaccharide O-antigen ABC transport system ATP-binding component LPS (VF0561)	20.187	3.846	0.000
	<i>cps</i>	glycosyltransferase Capsule (VF0881)	3.487	3.850	0.000
			3.715	3.926	0.000
	<i>SH_RS01805</i>	polysaccharide biosynthesis tyrosine autokinase Capsule (VF0003)	5.190	3.886	0.000
	<i>rfbB</i>	O-antigen export ABC transporter ATP-binding protein RfbB LPS (VF0561)	39.913	3.889	0.000
			20.196	3.833	0.000
	<i>SSP_RS00310</i>	polysaccharide biosynthesis tyrosine autokinase Capsule (VF0003)	5.420	3.921	0.000
	<i>ABTW07_0099</i>	UDP-glucose 4-epimerase Capsule (VF0465)	14.293	3.923	0.000
	<i>galE</i>	UDP-glucose 4-epimerase LOS (VF0044)	8.268	4.029	0.000
			17.993	4.272	0.000
			12.979	4.589	0.000
			7.627	3.875	0.000
			7.573	3.457	0.000
			9.829	2.376	0.021
		UDP-glucose 4-epimerase GalE O-antigen (VF0392)	11.919	3.947	0.000
		UDP-glucose 4-epimerase GalE Capsule (VF0465)	12.840	3.995	0.000
			13.873	4.002	0.000
			24.954	4.005	0.000
			13.313	4.011	0.000
			12.862	4.011	0.000
			13.910	4.057	0.000
			10.615	3.867	0.000
			10.494	3.863	0.000
			13.158	3.877	0.000
			11.912	3.774	0.000
			10.930	3.716	0.000
			6.625	2.879	0.001
			10.390	3.574	0.000
		UDP-glucose 4-epimerase GalE Polysaccharide capsule (VF0659)	11.051	2.276	0.035
			10.901	2.297	0.033
			11.036	2.305	0.033
			11.043	2.305	0.033
			11.043	2.305	0.033
			7.872	3.239	0.000
	<i>mrsA/glmM</i>	phosphoglucosamine mutase Exopolysaccharide (VF0755)	12.965	3.640	0.000
			17.921	4.001	0.000
			17.921	4.001	0.000
			17.274	3.864	0.000
			18.274	3.869	0.000
			18.286	3.902	0.000
	<i>SPD_RS01765</i>	oligosaccharide flippase family protein Capsule (VF0144)	18.231	3.902	0.000
	<i>kfiC</i>	lipopolysaccharide biosynthesis protein LOS (VF0044)	3.677	4.156	0.000
	<i>LPG_RS03830</i>	ATP-binding cassette domain-containing protein LPS (VF0171)	16.880	4.688	0.000
	<i>epsG</i>	glycosyltransferase Capsule (VF0881)	8.090	4.351	0.000
			42.708	4.472	0.000
	<i>SSA_RS07375</i>	ABC transporter ATP-binding protein Capsule (VF0144)	6.765	4.591	0.000
	<i>SAK_RS06315</i>	tyrosine-protein kinase Capsule (VF0274)	6.765	4.591	0.000
	<i>SSU05_0573</i>	Cps2J Capsule (VF0144)	14.597	4.625	0.000
	<i>GBS_RS06595</i>	tyrosine-protein kinase Capsule (VF0274)	12.903	4.641	0.000
	<i>C8J_RS06935</i>	glycosyltransferase family 2 protein Capsule (VF0323)	7.405	4.802	0.000
	<i>Cj1440c</i>	sugar transferase Capsule (VF0323)	24.996	4.831	0.000
	<i>SAG_RS08545</i>	glycosyltransferase Capsule (VF0274)	9.160	5.069	0.000
	<i>Cj1138</i>	glycosyltransferase LOS (VF0326)	10.392	5.407	0.000
			14.925	5.715	0.000
			13.268	5.883	0.000
Invasion (VFC0083)	<i>iap/cwhA</i>	P60 extracellular protein; invasion associated protein lap P60 (VF0068)	5.849	2.207	0.021
			5.899	2.262	0.018
			6.036	2.274	0.018
			6.036	2.274	0.018
			6.036	2.274	0.018
			6.078	2.274	0.018
			6.078	2.274	0.018
			6.078	2.274	0.018
			6.078	2.274	0.018
			6.078	2.274	0.018
			6.078	2.274	0.018
			11.567	2.395	0.025
			5.887	2.439	0.011
	<i>bca</i>	C protein alpha-antigen -C protein (VF0276)	2.423	2.513	0.000
	<i>aut</i>	autolysin Auto (VF0348)	5.813	2.629	0.007
	<i>tcyJ</i>	cystine ABC transporter substrate-binding protein Peritrichous flagella (VF1154)	4.103	2.225	0.013

Motility (VFC0204)	<i>flgJ</i>	peptidoglycan hydrolase Flagella (VF0519)	3.682	2.261	0.010
			3.877	2.660	0.002
			4.075	2.352	0.009
	<i>fliY</i>	cystine ABC transporter substrate-binding protein Peritrichous flagella (VF0967)	3.330	2.441	0.003
	<i>flmH</i>	short chain dehydrogenase/reductase family oxidoreductase Polar flagella (VF0473)	30.850	2.470	0.039
			6.475	2.097	0.025
	<i>fliI</i>	flagellum-specific ATP synthase FliI Flagella (VF0430)	2.780	2.032	0.007
		flagellum-specific ATP synthase Flagella (VF0519)	4.599	2.999	0.000
			4.364	3.040	0.000
			4.870	3.127	0.000
	<i>ASA_RS06915</i>	CobQ/CobB/MinD/ParA family protein Polar flagella (VF0473)	11.234	3.251	0.001
	<i>B565_RS05765</i>	CobQ/CobB/MinD/ParA family protein Polar flagella (VF0473)	11.234	3.251	0.001
	<i>AHML_RS07540</i>	CobQ/CobB/MinD/ParA family protein Polar flagella (VF0473)	22.001	3.345	0.002
	<i>GBAA_RS23245</i>	iron ABC transporter permease NEAT-type hemophore-mediated heme uptake system (VF0672)	7.191	2.003	0.042
	<i>dahP</i>	DAHPh synthetase Vanchrobactin (VF1246)	1.945	2.017	0.001
	<i>pdtrfI</i>	putative racemase Pyridine-2;6-dithiocarboxylic acid (PDTC) (VF0938)	3.435	2.056	0.014
	<i>irtB</i>	Iron-regulated transporter IrtB ABC transporter (VF0818)	6.742	2.244	0.026
			6.320	2.132	0.033
			9.565	2.218	0.034
			6.470	2.068	0.040
			6.530	2.098	0.038
			6.398	2.103	0.036
			8.047	2.376	0.021
			7.944	2.377	0.021
			9.748	2.727	0.009
			9.762	2.727	0.009
			9.677	2.742	0.008
			5.677	2.258	0.021
			7.005	2.263	0.026
			8.133	2.305	0.026
			7.238	2.311	0.022
			6.292	2.194	0.024
			7.344	2.357	0.015
			7.379	2.327	0.017
	<i>mgtC</i>	Possible Mg ²⁺ transport P-type ATPase C MgtC MgtC (VF0289)	4.579	2.110	0.027
			4.565	2.121	0.026
	<i>GBAA_RS23240</i>	ABC transporter ATP-binding protein NEAT-type hemophore-mediated heme uptake system (VF0672)	9.658	2.195	0.042
	<i>allR</i>	DNA-binding transcriptional repressor AllR Allantion utilization (VF0572)	1.941	2.215	0.001
	<i>dhbE</i>	2;3-dihydroxybenzoate adenylase DhbE Bacillibactin (VF0586)	4.467	2.224	0.012
	<i>ybtP</i>	yersiniabactin ABC transporter ATP-binding/permease protein YbtP Ybt (VF0564)	8.240	2.312	0.021
			8.407	2.334	0.020
			8.667	2.284	0.021
			21.710	2.413	0.033
		yersiniabactin ABC transporter ATP-binding/permease protein YbtP Yersiniabactin siderophore (VF1115)	11.681	2.422	0.020
			11.861	2.449	0.019
			11.961	2.470	0.018
			9.981	2.273	0.028
			10.734	2.227	0.033
			10.319	2.301	0.027
			10.877	2.304	0.027
		ABC transporter ATP-binding protein/permease Yersiniabactin (VF0923)	11.538	2.145	0.041
		yersiniabactin ABC transporter ATP-binding/permease protein YbtP Yersiniabactin (VF0136)	9.087	2.234	0.028
			8.824	2.332	0.021
			22.051	2.456	0.028
			7.915	2.195	0.026
			16.153	2.245	0.045
	<i>pvdE</i>	pyoverdine biosynthesis protein PvdE Pyoverdine (VF0094)	21.670	2.301	0.045
			21.301	2.302	0.047
			22.018	2.345	0.044
			18.686	2.462	0.031
			11.559	2.395	0.027
			22.382	2.480	0.034
			16.373	2.485	0.028
			5.632	3.614	0.000
			22.709	2.370	0.041
	<i>ybtQ</i>	yersiniabactin ABC transporter ATP-binding/permease protein YbtQ Ybt (VF0564)	19.056	2.275	0.045
	<i>hasD</i>	unknown protein HasA (VF1254)	10.034	2.294	0.030
	<i>narG</i>	nitrate reductase subunit alpha Nitrate reductase (VF0302)	7.829	2.247	0.033
			7.734	2.275	0.030

Nutritional/Metabolic
factor (VFC0272)

		6.769	2.114	0.037
<i>ctpV</i>	copper-translocating P-type ATPase Copper exporter (VF0849)	34.325	2.360	0.050
<i>PMI_RS12860</i>	ABC transporter ATP-binding protein/permease Yersiniabactin-related (VF1251)	19.782	2.375	0.035
<i>lysA</i>	Probable diaminopimelate decarboxylase Lysine synthesis (VF0815)	7.240	2.215	0.027
	diaminopimelate decarboxylase Lysine synthesis (VF0815)	7.042	2.073	0.037
		7.257	2.079	0.038
		7.265	2.079	0.038
		6.963	2.250	0.022
		6.502	2.223	0.026
		6.994	2.273	0.020
		5.783	2.063	0.030
		7.116	2.287	0.020
		6.138	2.180	0.021
		5.518	2.053	0.028
		6.168	2.189	0.020
		7.162	2.246	0.023
		7.159	2.288	0.020
		7.010	2.329	0.017
		5.709	2.600	0.006
		5.615	2.615	0.006
		6.851	2.454	0.012
		6.547	2.427	0.011
		6.895	2.394	0.015
		6.944	2.379	0.015
		6.967	2.379	0.016
		6.771	2.379	0.015
		7.148	2.239	0.024
<i>iroC</i>	ATP binding cassette transporter Salmochelin siderophore (VF0230)	20.724	2.382	0.041
		19.299	2.383	0.040
		20.194	2.407	0.039
		19.585	2.354	0.042
	ABC transporter Sal (VF0563)	74.732	4.374	0.000
		70.602	4.426	0.000
<i>iroB</i>	salmochelin/enterobactin export ABC transporter IroC Salmochelin synthesis and transport (VF0983)	19.964	2.386	0.040
<i>barA</i>	siderophore efflux system of the ABC superfamily Acinetobactin (VF0467)	23.338	2.371	0.045
		25.182	2.406	0.045
		25.208	2.409	0.045
		24.403	2.428	0.042
		24.279	2.429	0.043
<i>pchl</i>	ABC transporter ATP-binding protein Pyochelin (VF0095)	6.873	2.454	0.015
<i>PMI_RS01115</i>	ABC transporter ATP-binding protein Proteobactin (VF1252)	7.860	2.455	0.013
<i>barB</i>	siderophore efflux system of the ABC superfamily Acinetobactin (VF0467)	4.807	2.366	0.008
		9.156	2.574	0.012
		15.012	4.588	0.000
		14.748	4.569	0.000
		13.402	4.651	0.000
<i>ASA_RS16580</i>	heme ABC transporter ATP-binding protein Heme uptake system (VF0650)	7.966	2.513	0.013
<i>PSEEN_RS11625</i>	ABC transporter ATP-binding protein/permease Pseudomonine (VF0936)	21.810	2.597	0.024
<i>irp6C</i>	ABC transporter ATP-binding protein Siderophore-dependent iron uptake system (VF0738)	8.291	2.724	0.005
		10.034	3.014	0.003
<i>hmuV</i>	ABC transporter ATP-binding protein ABC-type heme transporter (VF0739)	10.835	2.744	0.009
<i>NWMN_RS12025</i>	iron chelate uptake ABC transporter family permease subunit Staphyloferrin A (VF1013)	7.814	2.765	0.003
<i>ctpC</i>	manganese-exporting P-type ATPase CtpC P(1)-type Mn ²⁺ transporting ATPase (VF0848)	16.525	2.804	0.013
		16.674	2.928	0.009
		18.076	2.932	0.009
		17.606	2.901	0.010
		17.646	3.036	0.006
		23.175	2.338	0.043
<i>fepC</i>	iron-enterobactin transporter ATP-binding protein Ent (VF0562)	14.863	2.926	0.008
<i>shuV</i>	ATP-binding component of hemin transport system Shu (VF0256)	16.638	2.993	0.006
<i>bauE</i>	ferric siderophore ABC transporter; ATP-binding protein BauE Acinetobactin (VF0467)	8.272	3.029	0.001
		18.101	3.181	0.003
		8.551	3.044	0.001
<i>ciuD</i>	ABC transporter ATP-binding protein Ciu iron uptake and siderophore biosynthesis system (VF0740)	18.398	3.070	0.005
		13.511	3.075	0.002
		13.511	3.075	0.002
		13.348	3.076	0.002
		13.511	3.078	0.002

<i>glnA1</i>	type I glutamate--ammonia ligase Glutamine synthesis (VF0816)	13.105	3.094	0.002
		13.526	3.105	0.002
		14.000	3.110	0.002
		14.000	3.110	0.002
		13.531	3.125	0.002
		13.607	3.158	0.001
		13.636	3.161	0.001
		13.463	3.162	0.001
		13.470	3.179	0.001
		13.593	3.198	0.001
		13.829	3.250	0.001
	Probable glutamine synthetase; type I (GlnA1) Glutamine synthesis (VF0816)	13.526	3.105	0.002
	glutamine synthetase Glutamine synthesis (VF0816)	13.809	3.128	0.002
		17.308	4.272	0.000
		9.353	4.121	0.000
<i>chuV</i>	ATP-binding hydrophilic protein ChuV Chu (VF0227)	18.806	3.250	0.002
<i>mgtB</i>	Mg2+ transport protein MgtBC (VF0106)	38.065	3.155	0.008
		40.561	3.277	0.006
		41.056	3.279	0.007
		41.220	3.292	0.006
		41.515	3.296	0.006
		40.919	3.299	0.006
		41.666	3.300	0.006
		42.793	3.313	0.006
		42.864	3.317	0.006
		41.573	3.317	0.006
		42.354	3.320	0.006
		85.552	3.334	0.008
		42.215	3.334	0.006
		58.121	3.453	0.005
<i>PFL_RS27280</i>	type I secretion system permease/ATPase HasA-type hemophore-mediated heme uptake system (VF0940)	6.162	3.362	0.000
<i>DDA3937_RS14675</i>	ABC transporter ATP-binding protein Chrysobactin (VF1242)	24.382	3.307	0.003
<i>fagC</i>	ABC transporter ATP-binding protein ABC transporter (VF0741)	15.697	3.395	0.001
<i>isdE</i>	iron-regulated surface determinant protein E Isd (VF0015)	4.734	2.258	0.018
	heme ABC transporter substrate-binding protein IsdE NEAT-type surface protein; direct heme uptake system (VF1106)	5.511	3.437	0.000
<i>feoB</i>	Fe(2+ FupA (VF0550)	5.353	3.320	0.000
		5.698	3.435	0.000
		5.858	3.469	0.000
		5.858	3.469	0.000
		10.672	3.584	0.000
		5.858	3.469	0.000
		5.858	3.469	0.000
		5.894	3.478	0.000
	ferrous iron transporter B FeoAB (VF0160)	6.956	3.673	0.000
		13.287	3.819	0.000
		7.670	3.824	0.000
		6.978	3.680	0.000
		7.006	3.631	0.000
<i>pchE</i>	dihydroaeruginosic acid synthetase PchE Pyochelin (VF0095)	5.724	3.575	0.000
<i>YPO_RS20735</i>	type I secretion system permease/ATPase HasA-type hemophore-mediated heme uptake system (VF1029)	13.712	3.640	0.000
<i>DDA3937_RS07730</i>	DHA2 family efflux MFS transporter permease subunit Achromobactin (VF1243)	11.393	3.724	0.000
<i>allB</i>	allantoinase Allantoin utilization (VF0572)	12.376	3.778	0.000
<i>oppA</i>	hypothetical protein Oligopeptide-binding protein oppA (VF0786)	15.981	3.868	0.000
	peptide ABC transporter substrate-binding protein Oligopeptide-binding protein oppA (VF0786)	15.981	3.868	0.000
		15.981	3.868	0.000
		15.952	3.868	0.000
		15.865	3.879	0.000
		15.902	3.918	0.000
		17.784	3.989	0.000
		18.178	4.045	0.000
		16.137	3.895	0.000
<i>exiT</i>	ABC transporter ATP-binding protein/permease Exochelin (smegmatis) (VF0819)	12.193	3.905	0.000
<i>isdF</i>	iron-regulated surface determinant protein F; ATP-binding-cassette-type transmembrane transporter Isd (VF0015)	10.164	4.191	0.000
		7.571	4.196	0.000
		7.606	4.202	0.000
		7.592	4.202	0.000
		7.635	4.209	0.000
		7.474	4.263	0.000
		7.471	4.305	0.000

	<i>irp2</i>	yersiniabactin non-ribosomal peptide synthetase HMWP2 Yersiniabactin siderophore (VF1115)	7.635	4.209	0.000
			7.578	4.228	0.000
			7.578	4.228	0.000
			7.452	4.258	0.000
			7.169	4.432	0.000
			7.343	4.443	0.000
			7.230	4.464	0.000
			7.155	4.464	0.000
			7.155	4.464	0.000
			7.163	4.464	0.000
			7.266	4.491	0.000
			7.169	4.501	0.000
		yersiniabactin non-ribosomal peptide synthetase HMWP2 Ybt (VF0564)	14.190	4.581	0.000
		yersiniabactin biosynthetic protein Irp2 Yersiniabactin (VF0136)	14.024	4.597	0.000
		amino acid adenylation domain-containing protein Yersiniabactin siderophore (VF1115)	7.479	4.231	0.000
	<i>irtA</i>	Iron-regulated transporter IrtA ABC transporter (VF0818)	16.840	2.185	0.050
			15.191	2.488	0.021
			12.873	2.391	0.023
			20.293	4.234	0.000
			18.411	4.099	0.000
			13.750	3.930	0.000
			20.006	4.271	0.000
			19.044	4.279	0.000
			15.739	4.286	0.000
			33.225	4.508	0.000
			26.630	4.538	0.000
			18.665	2.340	0.038
			9.134	2.307	0.021
			7.018	2.032	0.035
			8.993	2.211	0.027
	<i>vctC</i>	iron chelate ABC transporter ATP-binding protein VctC VctPDGC system (VF0638)	13.368	4.556	0.000
			5.955	2.603	0.007
			5.904	2.615	0.007
Post-translational modification (VFC0315)	<i>lgt</i>	prolipoprotein diacylglycerol transferase Lipoprotein diacylglycerol transferase lgt (VF0790)	8.673	2.220	0.033
			8.673	2.220	0.033
			8.494	2.271	0.029
			8.465	2.277	0.028
			8.465	2.277	0.028
			8.465	2.277	0.028
			8.465	2.277	0.028
			8.465	2.277	0.028
	<i>dltA</i>	D-alanine--poly(phosphoribitol D-alanine-polyphosphoribitol ligase DltA (VF0795)	12.954	2.293	0.045
			12.954	2.293	0.045
			12.960	2.293	0.045
			12.968	2.293	0.045
			12.968	2.293	0.045
			10.819	2.272	0.039
			12.956	2.272	0.047
			13.044	2.272	0.047
			12.974	2.293	0.045
			12.974	2.293	0.045
	<i>lspA</i>	signal peptidase II Lsp (VF0351)	3.844	2.309	0.011
			3.859	2.309	0.011
			3.859	2.309	0.011
			4.006	2.370	0.009
			4.006	2.370	0.009
	<i>bvgA</i>	response regulator transcription factor BvgAS (VF0336)	5.388	2.010	0.036
	<i>relA</i>	Probable GTP pyrophosphokinase RelA (ATP:GTP 3'- pyrophosphotransferase RelA (VF0287)	71.379	3.911	0.001
		GTP diphosphokinase RelA (VF0260)	6.290	2.318	0.014
			6.247	2.270	0.016
			6.190	2.270	0.016
	<i>bvrR</i>	response regulator transcription factor BvrR-BvrS (VF0368)	15.809	2.379	0.034
			3.220	2.416	0.006
			3.220	2.416	0.006
			5.436	2.484	0.010
			3.486	2.504	0.004
			3.471	2.544	0.004
			5.224	2.583	0.006
			6.620	2.593	0.011
			6.179	2.600	0.010
			6.799	2.645	0.009

Regulation (VFC0301)	<i>phoR</i>	Possible two component system response sensor kinase membrane associated PhoR PhoP (VF0286)	9.876	2.703	0.012
			8.577	3.498	0.000
			8.606	3.655	0.000
			8.861	3.884	0.000
			5.185	2.312	0.015
			5.200	2.322	0.015
			5.200	2.322	0.015
			3.889	2.033	0.023
			4.467	2.119	0.021
			4.854	2.132	0.023
			4.166	2.136	0.017
			4.278	2.144	0.018
			4.290	2.147	0.020
			4.926	2.182	0.020
			4.969	2.194	0.020
	<i>agrA</i>	response regulator AgrA/AgrC (VF0793)	2.516	3.174	0.000
		LytTR family DNA-binding domain-containing protein AgrA/AgrC (VF0793)	2.516	3.174	0.000
			2.516	3.174	0.000
			2.516	3.174	0.000
			2.516	3.174	0.000
	<i>senX3</i>	histidine kinase SenX3 (VF0857)	13.588	3.178	0.003
			13.417	3.168	0.003
			12.578	3.652	0.000
		sensor histidine kinase SenX3 (VF0857)	14.896	3.277	0.002
			13.488	3.455	0.001
			13.753	3.644	0.000
		two-component sensor histidine kinase SenX3 (VF0857)	13.601	3.560	0.001
			12.858	2.887	0.008
			12.780	3.027	0.004
			14.508	3.384	0.001
			12.765	3.027	0.004
			12.898	3.640	0.000
			9.116	2.494	0.019
			13.847	3.604	0.001
		two-component system sensor histidine kinase SenX3 SenX3 (VF0857)	11.619	3.457	0.001
			15.272	3.327	0.002
			13.409	3.583	0.001
		two component sensor histidine kinase SenX3 SenX3 (VF0857)	11.192	3.419	0.001
			13.354	3.558	0.001
		Sensor-like histidine kinase senX3 SenX3 (VF0857)	12.729	2.856	0.009
	<i>lisk</i>	HAMP domain-containing protein LisR/LisK (VF0792)	19.285	3.552	0.001
			17.629	3.577	0.001
			21.687	3.646	0.001
			21.730	3.646	0.001
			21.764	3.629	0.001
			21.910	3.629	0.001
		two-component sensor histidine kinase LisR/LisK (VF0792)	21.953	3.634	0.001
	<i>phoP</i>	Possible two component system response transcriptional positive regulator PhoP PhoP (VF0286)	17.815	2.283	0.041
			21.004	4.085	0.000
			20.192	4.090	0.000
			23.169	4.217	0.000
			23.062	4.220	0.000
			20.840	4.250	0.000
			23.845	4.262	0.000
			23.545	4.289	0.000
			24.765	4.319	0.000
			22.447	4.326	0.000
			23.827	4.340	0.000
			24.305	4.344	0.000
			48.623	4.348	0.000
			23.325	4.351	0.000
			22.544	4.366	0.000
			25.090	4.398	0.000
			24.799	4.437	0.000
	<i>mntA</i>	ABC transporter ATP-binding protein MntA MntABC (VF0455)	4.370	2.003	0.030
			3.964	2.030	0.021
			3.800	2.083	0.018
			4.784	2.171	0.021
			7.660	2.195	0.033
	<i>tig/ropA</i>	trigger factor Trigger factor (VF1056)	5.040	2.206	0.020
			13.225	2.241	0.046
			13.510	2.243	0.046
			13.525	2.243	0.046
			13.583	2.248	0.045
			13.583	2.248	0.045
			12.453	2.288	0.040

Stress survival (VFC0282)			12.453	2.288	0.040
	<i>ureB</i>	urease beta subunit UreB; urea amidohydrolase Urease (VF0050)	8.981	2.257	0.039
			8.981	2.257	0.039
			8.981	2.257	0.039
			8.666	2.261	0.038
			7.818	2.288	0.033
	<i>mpa</i>	Putative AAA-family ATPase Proteasome-associated proteins (VF0847)	19.059	3.473	0.001
		proteasome ATPase Proteasome-associated proteins (VF0847)	19.059	3.473	0.001
			20.279	3.494	0.001
			20.279	3.494	0.001
			20.279	3.494	0.001
			20.750	3.502	0.001
			20.643	3.518	0.001
			20.784	3.530	0.001
			20.720	3.531	0.001
			20.742	3.534	0.001
			20.514	3.587	0.000
			20.514	3.587	0.000
			20.449	3.589	0.000
			20.629	3.597	0.000
			20.628	3.597	0.000
			20.577	3.604	0.000
			20.607	3.605	0.000
			20.607	3.605	0.000
		proteasome-associated ATPase Proteasome-associated proteins (VF0847)	20.629	3.597	0.000
		putative ATPase Proteasome-associated proteins (VF0847)	12.933	3.697	0.000
	<i>bsh</i>	bile salt hydrolase BSH (VF0350)	5.808	4.023	0.000
			5.808	4.023	0.000
			5.808	4.023	0.000
			5.808	4.023	0.000
			5.808	4.023	0.000
			5.822	4.031	0.000
			5.822	4.031	0.000
			5.937	4.063	0.000
			6.266	4.178	0.000
	<i>msrA/B(pilB)</i>	trifunctional thioredoxin/methionine sulfoxide reductase A/B protein MsrAB (VF0456)	9.463	4.634	0.000
			10.323	4.316	0.000
			10.323	4.316	0.000
			10.323	4.316	0.000
			10.323	4.316	0.000
			10.323	4.316	0.000
			10.359	4.323	0.000
			10.294	4.335	0.000
			20.004	4.476	0.000
Others (VFC0346)	<i>cmaA2</i>	class I SAM-dependent methyltransferase Mycolic acid trans-cyclopropane synthetase (VF0802)	2.158	2.116	0.001
			2.172	2.141	0.001
			2.115	2.160	0.001
			1.912	2.162	0.001
			1.912	2.162	0.001
	<i>mmaA4</i>	hydroxymycolate synthase MmaA4 MmaA4 (VF0314)	2.472	3.528	0.000
			2.186	2.141	0.001
	<i>secA2</i>	preprotein translocase subunit SecA Accessory secretion factor (VF0808)	3.396	3.806	0.000
			20.884	3.927	0.000
		accessory Sec system translocase SecA2 Accessory secretion factor (VF0808)	23.496	3.872	0.000
			23.991	3.907	0.000
			23.653	3.925	0.000
			25.170	3.928	0.000
			25.039	3.929	0.000
			25.405	3.934	0.000
			25.250	3.948	0.000
			23.350	3.954	0.000
			26.199	4.029	0.000
			20.331	4.048	0.000
			17.952	4.059	0.000
			28.067	4.060	0.000
			28.089	4.060	0.000
			31.355	4.061	0.000
			28.218	4.063	0.000
			27.817	4.065	0.000
			27.154	4.070	0.000
			29.690	4.088	0.000
			29.690	4.088	0.000

			29.690	4.088	0.000
			30.832	4.110	0.000
			29.687	4.115	0.000
			31.365	4.120	0.000
			27.179	4.122	0.000
			25.900	4.187	0.000
			19.033	4.297	0.000
			21.698	4.342	0.000
			21.044	4.385	0.000
Species-associated genes	esp	Enterococcal surface protein (Enterococcus faecalis)	5.273	2.033	0.035
			5.527	3.740	0.000

Annex 5. Unique virulence factors in the clinical miscarriage (CM) group.

Category	VF gene name	Description
Adherence (VFC0001)	AAX78184	LdaA Locus for diffuse adherence (Lda), afimbrial adhesin (VF1148)
	afnB	usher protein precursor Adhesive factor/rabbit 1 (AF/R1) pili (VF1146)
	clfA	Clumping factor A, fibrinogen-binding protein Clumping factor (VF0004)
	cswD	CS12 fimbria outer membrane usher protein precursor Adhesive fimbriae (VF0213)
	DIP_RS23695	CshA/CshB family fibrillar adhesin-related protein DIP2093 (VF0604)
	ebp	cell surface elastin binding protein EbpS (VF0008)
	ecbA/fss3	cell surface elastin binding protein Autolysin alt (VF0997)
	ecbA/fss3	Collagen binding MSCRAMM, EcbA/Fss3 EcbA (VF0539)
	ehaB	autotransporter adhesin EhaB EhaB, AIDA-I type (VF1128)
	epf	extracellular protein Extracellular factor (VF1057)
	fasD	FasD Adhesive fimbriae (VF0213)
	fedC	FedC protein Adhesive fimbriae (VF0213)
	fimC	Chaperone protein fimC precursor Type 1 fimbriae (VF0221)
	fimC	periplasmic chaperone Type I fimbriae (VF0566)
	fimG	FimG protein precursor Type 1 fimbriae (VF0221)
	fimH	FimH protein precursor Type 1 fimbriae (VF0221)
	fimH	type 1 fimbrial adhesin precursor Type I fimbriae (VF0566)
	fimK	transcriptional regulator Type I fimbriae (VF0566)
	fimU	type 4 fimbrial biogenesis protein FimU Type IV pili (VF0082)
	flp1	Flp family type IVb pilin The tad locus (VF0754)
	flp2	Flp family type IVb pilin The tad locus (VF0754)
	focC	F1C periplasmic chaperone F1C fimbriae (VF0224)
	ftoD	FotD Adhesive fimbriae (VF0213)
	inlF	internalin F InlF (VF0447)
	lcl	eukaryotic-like collagen-like protein Lpg2644 (VF0800)
	legA6	Dot/Icm type IV secretion system effector LegA6 Dot/Icm T4SS secreted effectors (VF0798)
	mptC	alpha-(1-2)-mannopyranosyltransferase MptC CdiLAM (VF0606)
	mrkB	MrkB Type 3 fimbriae (VF1208)
	mrp	hypothetical protein Muramidase-released protein (VF1052)
	mshA	prepilin-type N-terminal cleavage/methylation domain-containing protein MSHA type IV pili (VF0477)
	mshM	MSHA biogenesis protein MshM MSHA pili (VF0515)
	PD_RS00110	Tfp pilus assembly protein FimT/FimU Type IV pili (VF1240)
	PD_RS08895	pilus assembly protein PilP Type IV pili (VF1240)
	PD_RS10195	pilin Type IV pili (VF1240)
	pilA	type 4 fimbrial precursor PilA Type IV pili (VF0082)
	pilC	type IV pilus assembly protein PilC Type IV pili (VF0431)
	pilE	type 4 fimbrial biogenesis protein PilE Type IV pili (VF0082)
	pilE4	prepilin-type N-terminal cleavage/methylation domain-containing protein Type IV pili (VF0458)
	pilF	Type IV pilus assembly protein PilF Type IV pili (VF0458)
	pilF	type 4 fimbrial biogenesis protein PilF Type IV pili (VF0082)
	pilH	pilin-like protein may involving in pseudopilus formation Type IV pili (VF0075)
	pilM	pilus assembly protein PilM Type IV pili (VF0731)
	pilN	type IV pilus inner membrane platform protein PilN Type IV pili (VF0082)
	pilO	type IV pilus inner membrane platform protein PilO Type IV pili (VF0075)
	pilP	type IV pilus biogenesis protein PilP Type IV pili (VF0082)
	pilZ	PilZ domain-containing protein Type IV pili (VF0082)
	pilZ	type IV pilus assembly protein PilZ Type IV pili (VF0075)
	pilZ	type IV pilus biogenesis protein PilZ Type IV pili (VF0155)

	<i>pmpD</i>	polymorphic membrane protein PmpD Pmps (VF1346)
	<i>psaB</i>	chaperone protein PsaB Myf/pH6 antigen (VF0134)
	<i>rrgB</i>	cell wall surface anchor family protein, major pilin RrgB RlrA islet (VF0529)
	<i>rrgC</i>	cell wall surface anchor family protein, minor pili RrgC RlrA islet (VF0529)
	<i>safB</i>	putative fimbriae assembly chaparone Saf (VF0951)
		pili assembly chaperone SafB Saf (VF0951)
	<i>sefC</i>	SEF14/SEF18 fimbria usher protein SefC Sef (VF0952)
	<i>sfaE</i>	S fimbriae periplasmic chaperone SfaE S fimbriae (VF0222)
	<i>SGO_RS09810</i>	LPXTG-anchored isopeptide-forming adhesin PadA PadA SGO_2005 (VF1086)
	<i>spaE</i>	SpaD-type pili minor subunit SpaE SpaD-type pili (VF0726)
	<i>staD</i>	fimbrial protein Sta (VF0953)
	<i>stbC</i>	fimbrial outer membrane usher protein Stb (VF0954)
		fimbrial assembly protein Stb (VF0954)
	<i>sthB</i>	fimbrial biogenesis chaperone SthA Sth (VF0960)
	<i>stiB</i>	long polar fimbrial chaperone LpfB Sti (VF0961)
		fimbrial biogenesis chaperone StiB Sti (VF0961)
	<i>stkD</i>	fimbrial protein Stk (VF0963)
	<i>tadA</i>	Flp pilus assembly complex ATPase component TadA Colonization factor Citrobacter (CFC) type IV fimbriae (VF1209)
	<i>tadG</i>	pilus assembly protein The tad locus (VF0754)
	<i>tcfD</i>	fimbrial protein TcfD Tcf (VF0964)
	<i>tibA</i>	autotransporter adhesin/invasin glycoprotein TibA AIDA-I type (VF1132)
	<i>VV1_RS01705</i>	GspH/FimT family pseudopilin ChiRP (VF0609)
	<i>yagZ/ecpA</i>	E. coli common pilus structural subunit EcpA ECP (VF0404)
Biofilm (VFC0271)	<i>alg44</i>	alginate biosynthesis protein Alg8 Alginate (VF0091)
	<i>algK</i>	alginate biosynthesis protein Alg44 Alginate (VF0091)
	<i>csuD</i>	Csu pilus usher protein CsuD Csu fimbriae (VF0461)
	<i>cupB4</i>	chaperone CupB4 CupB fimbriae (VF0925)
	<i>cupC3</i>	usher CupC3 CupC fimbriae (VF0926)
	<i>mrkB</i>	fimbrial chaperone protein mrkB precursor Type 3 fimbriae (VF0567)
Effector delivery system (VFC0086)	<i>aaiE</i>	type VI secretion system baseplate subunit TssF AAI/SCI-II T6SS (VF1121)
	<i>aec32</i>	Hcp family type VI secretion system effector ACE T6SS (VF1122)
	<i>APECO1_RS15990</i>	Hcp family type VI secretion system effector SCI-I T6SS (VF1120)
	<i>app</i>	adhesion and penetration autotransporter App Adhesion and penetration protein (VF0893)
	<i>ATU_RS20350</i>	Hcp family type VI secretion system effector Imp (VF1258)
	<i>C_RS16090</i>	Hcp family type VI secretion system effector SCI-I T6SS (VF1120)
	<i>cem1</i>	Coxiella Dot/lcm type IVB secretion system translocated effector, DnaJ domain-containing protein T4SS secreted effectors (VF0696)
	<i>dotB</i>	Dot/lcm type IV secretion system ATPase DotB Dot/lcm (VF0156)
	<i>dotU</i>	type VI secretion system protein TssL, short form T6SS-III (VF0784)
	<i>EMY_RS24785</i>	EscR/YscR/HrcR family type III secretion system export apparatus protein PAI-3 encoded (VF1272)
	<i>EC55989_RS17090</i>	type VI secretion protein Vask SCI-I T6SS (VF1120)
	<i>EC958_RS15770</i>	type VI secretion system ImpA family N-terminal domain-containing protein SCI-I T6SS (VF1120)
	<i>ECABU_RS15695</i>	type VI secretion system ImpA family N-terminal domain-containing protein SCI-I T6SS (VF1120)
	<i>eccC3</i>	Type VII secretion system protein EccC3 ESX-3 (VF0521)
	<i>ECE24377A_RS17310</i>	type VI secretion system ImpA family N-terminal domain-containing protein SCI-I T6SS (VF1120)
	<i>ECH_RS04235</i>	hypothetical protein T4SS (VF1269)
	<i>ECNA114_RS14750</i>	type VI secretion system contractile sheath small subunit SCI-I T6SS (VF1120)
	<i>ECNA114_RS14850</i>	type VI secretion system baseplate subunit TssE SCI-I T6SS (VF1120)
	<i>ECO103_RS17720</i>	Hcp family type VI secretion system effector SCI-I T6SS (VF1120)
	<i>ECO103_RS17790</i>	type VI secretion system ImpA family N-terminal domain-containing protein SCI-I T6SS (VF1120)
	<i>ECOK1_RS16710</i>	type VI secretion system ImpA family N-terminal domain-containing protein SCI-I T6SS (VF1120)
	<i>epeA</i>	serine protease autotransporter EpeA EpeA (VF1166)
	<i>epsK</i>	type II secretion system minor pseudopilin GspK Eps T2SS (VF0613)
	<i>esaG</i>	TIGR01741 family protein Type VII secretion system (VF0403)
		antitoxin YezG family protein Type VII secretion system (VF0403)

<i>escR</i>	T3SS structure protein EscR ETT2 (VF1161)
<i>espC</i>	Type III secretion system secreted protein, serine protease EspC (VF0173)
<i>espP</i>	autotransporter, serine protease EspP (VF0208)
<i>gspJ</i>	general secretion pathway protein J T2SS (VF1335)
<i>gspK</i>	type II secretion system minor pseudopilin GspK Etp (VF1178)
<i>hcp</i>	type VI secretion system effector Hcp SCI-I T6SS (VF1120)
	type VI secretion system tube protein Hcp CTS2 (VF1264)
<i>hcp/tssD</i>	type VI secretion system protein, Hcp family T6SS (VF0569)
<i>hilC</i>	AraC family transcriptional regulator TTSS (SPI-1 encode) (VF0116)
	invasion transcriptional regulator HilC TTSS (SPI-1 encode) (VF0116)
<i>hopI1</i>	DnaJ domain-containing protein P. syringae TTSS effectors (VF0911)
<i>hrcU</i>	type III secretion system export apparatus subunit SctU P. syringae TTSS (VF0909)
<i>icmF</i>	type VI secretion protein VasK T6SS-III (VF0784)
<i>icmW</i>	Dot/Icm type IV secretion system chaperone IcmW Dot/Icm (VF0156)
<i>iga</i>	IgA-specific serine endopeptidase IgA1 protease (VF0080)
<i>impA</i>	type VI secretion system ImpA family N-terminal domain-containing protein T6SS-III (VF0784)
<i>impF</i>	type VI secretion system baseplate subunit TssE T6SS-II (VF0783)
	type VI secretion system baseplate subunit TssE T6SS-III (VF0784)
<i>KPN2242_RS27900</i>	peptidoglycan DD-metalloendopeptidase family protein T6SS-III (VF0784)
<i>LPG_RS06270</i>	carbon storage regulator Lvh (Legionella vir homologs) type IVA secretion system (VF0799)
<i>MAFF_RS25875</i>	EscS/YscS/HrcS family type III secretion system export apparatus protein T3SS (VF1274)
<i>NGR_RS11020</i>	type III secretion protein T3SS2 (VF1299)
<i>PAU_RS17165</i>	winged helix-turn-helix domain-containing protein T3SS-2 (VF1287)
<i>pet</i>	Pet serine protease precursor, autotransporter Pet (VF0217)
<i>pscS</i>	type III secretion system protein PscS TTSS (VF0083)
<i>sctC</i>	type III secretion system outer membrane ring subunit SctC T3SS (VF1301)
<i>sctS</i>	type III secretion system export apparatus subunit SctS T3SS (VF1266)
	type III secretion system export apparatus subunit SctS T3SS-1 (VF1286)
<i>sctU</i>	type III secretion system export apparatus subunit SctU T3SS (VF1295)
<i>SG_RS24060</i>	hypothetical protein T6SS (VF0975)
<i>spa40</i>	type III secretion system export apparatus switch protein Spa40 TTSS (VF0118)
<i>spaP</i>	type III secretion system minor export apparatus protein SpaP TTSS (SPI-1 encode) (VF0116)
<i>spaS</i>	type III secretion system export apparatus switch protein SpaS TTSS (SPI-1 encode) (VF0116)
<i>ssaC</i>	type III secretion system secretin SsaC TTSS (SPI-2 encode) (VF0321)
<i>ssaU</i>	type III secretion system export apparatus switch protein SsaU TTSS (SPI-2 encode) (VF0321)
<i>STM0281</i>	type VI secretion system-associated protein SCI (Salmonella centrisome island) (VF0974)
<i>tke6</i>	PAAR domain-containing protein, putative T6SS effector Putida-T6SS secreted effectors (VF1354)
<i>tssB</i>	type VI secretion system contractile sheath small subunit SCI-I T6SS (VF1120)
	type VI secretion system contractile sheath small subunit T6SS (VF1337)
	type VI secretion system contractile sheath small subunit T6SS-A (VF1036)
<i>tssE</i>	type VI secretion system baseplate subunit TssE SCI-I T6SS (VF1120)
<i>tssF</i>	type VI secretion system baseplate subunit TssF CTS2 (VF1264)
	type VI secretion system baseplate subunit TssF T6SS (VF1337)
<i>tssK</i>	type VI secretion system baseplate subunit TssK HSI-2 (VF0943)
<i>tssL</i>	type VI secretion system protein TssL, short form SCI-I T6SS (VF1120)
<i>tssL2/icmH</i>	type IV secretion system protein IcmH/DotU Putida K2-T6SS (VF1352)
<i>tssL3/icmH</i>	type IVB secretion system protein IcmH/DotU Putida K3-T6SS (VF1353)
<i>tssM-5</i>	Type VI secretion system protein TssM T6SS-1 (VF0429)

	<i>vasJ</i>	type VI secretion system protein VasJ VAS T6SS (VF0335)
	<i>vat</i>	Haemoglobin protease Vacuolating autotransporter gene Vat (VF1112)
	<i>vgrG</i>	type VI secretion system tip protein VgrG T6SS-III (VF0784)
	<i>vipA/tssB</i>	type VI secretion system contractile sheath small subunit VipA T6SS (VF0569)
	<i>XCC_RS03430</i>	GspH/FimT family pseudopilin Xps (VF1304)
	<i>XCC_RS03445</i>	type II secretion system protein K Xps (VF1304)
	<i>xcpW</i>	general secretion pathway protein J Xcp secretion system (VF0084)
	<i>XFASM12_RS04000</i>	Tfp pilus assembly protein FimT/FimU Gsp (VF1307)
	<i>ysaR</i>	EscR/YscR/HrcR family type III secretion system export apparatus protein Ysa TTSS (VF1025)
	<i>yscS</i>	type III secretion system minor export apparatus protein YscS TTSS (VF0140)
	<i>ysrR</i>	response regulator transcription factor Ysa TTSS (VF1025)
	<i>yst1F</i>	type II secretion system F family protein Yst1 T2SS (VF0635)
Exoenzyme (VFC0251)	<i>gtfG</i>	KxYKxGKxW signal peptide domain-containing protein Streptococcal glucosyltransferases (VF1051)
	<i>nanH</i>	sialidase Sialidase (VF0391)
	<i>nuc</i>	thermonuclease family protein Nuclease (VF0880)
	<i>sepA</i>	extracellular serine protease SepA Serine protease (VF0982)
		serine protease autotransporter toxin SepA Serine protease (VF0982)
	<i>sigA</i>	serine protease, autotransporter SigA SigA (VF0119)
Exotoxin (VFC0235)	<i>cesP</i>	4'-phosphopantetheinyl transferase, CesP Cereulide (VF0583)
	<i>clbA</i>	colibactin biosynthesis phosphopantetheinyl transferase ClbA Colibactin (VF0573)
		colibactin biosynthesis phosphopantetheinyl transferase ClbA Colibactin (VF1179)
	<i>cry42Aa1</i>	Cry protein Insecticidal crystalline toxins (VF0658)
	<i>PSPPH_RS21800</i>	hypothetical protein Phytotoxin phaseolotoxin (VF0917)
	<i>vip2Ac</i>	Vip2Ac Vegetative insecticidal protein (VF0673)
Immune modulation (VFC0258)	<i>ABD1_RS00315</i>	NAD-dependent epimerase/dehydratase family protein Capsule (VF0465)
	<i>ABZJ_RS06290</i>	
	<i>BPSL2786</i>	acetyltransferase Capsule I (VF0436)
	<i>C8J_RS05590</i>	FdtA/QdtA family cupin domain-containing protein LOS (VF0326)
	<i>C8J_RS08645</i>	adenylyl-sulfate kinase Capsule (VF0323)
	<i>caf1A</i>	F1 capsule anchoring protein F1 antigen (VF0138)
	<i>caf1R</i>	F1 operon positive regulatory protein F1 antigen (VF0138)
	<i>cap8F</i>	type 8 capsular polysaccharide synthesis protein Cap8F Capsule (VF0003)
	<i>capB</i>	CapB, involved in Poly-gamma-glutamate synthesis Capsule (VF0141)
	<i>capC</i>	CapC, involved in Poly-gamma-glutamate synthesis Capsule (VF0141)
	<i>CFF8240_RS06845</i>	glycosyltransferase family 2 protein LOS (VF0326)
	<i>cps4B</i>	capsular polysaccharide biosynthesis protein Cps4B Capsule (VF0144)
	<i>cps4K</i>	capsular polysaccharide biosynthesis protein Cps4K Capsule (VF0144)
	<i>flmK</i>	glycosyltransferase family 39 protein LPS (VF0542)
	<i>lgtC</i>	glycosyltransferase LOS (VF0044)
		glycosyltransferase family 8 protein LOS (VF0078)
	<i>LPG_RS03810</i>	class I SAM-dependent methyltransferase LPS (VF0171)
	<i>lpxB</i>	lipid A disaccharide synthetase LPS (VF0056)
	<i>lpxC</i>	UDP-3-O-acyl-N-acetylglucosamine deacetylase LPS (VF0370)
	<i>lpxE</i>	phosphatase PAP2 family protei LPS (VF0542)
	<i>M3Q_RS01525</i>	LTA synthase family protein Capsule (VF0465)
	<i>mce2F</i>	MCE family protein Mce2 (VF0825)
		virulence factor mce family protein Mce2 (VF0825)
		MCE-family protein mce2F Mce2 (VF0825)
	<i>mce4A</i>	MCE family protein Mce4 (VF0827)
		virulence factor Mce family protein Mce4 (VF0827)
		virulence factor Mce4A Mce4 (VF0827)
		Mce family protein Mce4A Mce4 (VF0827)
		MCE family protein Mce4 (VF0827)
	<i>mce4B</i>	Mce family protein Mce4B Mce4 (VF0827)
		virulence factor Mce family protein Mce4 (VF0827)
		Hypothetical MCE-family protein Mce4 (VF0827)
	<i>mce4D</i>	MCE family protein Mce4 (VF0827)
		virulence factor Mce family protein Mce4 (VF0827)
	<i>mce4F</i>	hypothetical protein Mce4 (VF0827)
		hypothetical protein Mce4 (VF0827)
	<i>napA</i>	neutrophil activating protein NapA HP-NAP (VF0052)
	<i>PLES_RS09930</i>	SDR family oxidoreductase LPS (VF0085)
	<i>ptpA</i>	protein tyrosine phosphatase Tyrosine phosphatase (VF0852)
	<i>rhlA</i>	rhamnosyltransferase chain A Rhamnolipid (VF0089)

	<i>Rv0926</i>	dihydrodipicolinate reductase GPL locus (VF0841)
	<i>SH_RS01810</i>	capsular polysaccharide biosynthesis protein Cap8C Capsule (VF0003)
	<i>SH_RS01825</i>	NAD-dependent epimerase/dehydratase family protein Capsule (VF0003)
	<i>SH_RS01840</i>	hypothetical protein Capsule (VF0003)
	<i>SPCG_RS01810</i>	hypothetical protein Capsule (VF0144)
	<i>SSA_RS10870</i>	NAD-dependent epimerase/dehydratase family protein Capsule (VF0144)
	<i>STR_RS05200</i>	lipopolysaccharide biosynthesis protein Capsule (VF0144)
	<i>waaQ</i>	heptosyltransferase-like protein LOS (VF0044)
	<i>wbtT</i>	NAD-dependent epimerase/dehydratase family protein Capsular polysaccharide (VF0624)
	<i>wbtG</i>	glycosyl transferase family protein LPS (VF0542)
	<i>wbtN</i>	glycosyltransferase LPS (VF0542)
	<i>wzm</i>	O-antigen export system permease protein LPS (VF0367)
Invasion (VFC0083)	<i>ail</i>	attachment invasion locus protein Ail (VF0132)
	<i>ompB</i>	outer membrane protein OmpB ROmpB/Sca5 (VF0493)
Motility (VFC0204)	<i>BCE_RS08820</i>	flagellar hook protein FlgE Peritrichous flagella (VF0670)
	<i>flgC</i>	flagellar basal-body rod protein FlgC Flagella (VF0051)
		flagellar basal body rod protein FlgC Peritrichous flagella (VF0732)
	<i>flgK</i>	flagellar hook-associated protein 1 FlgK Flagella (VF0273)
		flagellar hook-associated protein 1 FlgK Flagella (VF0157)
	<i>flgL</i>	flagellar hook-associated protein FlgL Flagella (VF0685)
	<i>fliE</i>	flagellar hook-basal body complex protein FliE Flagella (VF0051)
		flagellar hook-basal body complex protein FliE Flagella (VF0114)
		flagellar hook-basal body complex protein FliE Peritrichous flagella (VF0732)
	<i>fliEL</i>	FliEL Lateral flagella (VF0474)
	<i>fliH</i>	flagellar assembly protein FliH Peritrichous flagella (VF0967)
		flagellar assembly protein H Flagella (VF0394)
	<i>fliJ</i>	flagellar biosynthesis chaperone FliJ Peritrichous flagella (VF0967)
		flagella biosynthesis chaperone FliJ Peritrichous flagella (VF1154)
		flagellar protein FliJ Flagella (VF0394)
	<i>fliK</i>	flagellar hook-length control protein FliK, putative Flagella (VF0519)
		flagellar hook-length control protein FliK Flagella (VF0157)
	<i>fliL</i>	flagellar basal body protein FliL Flagella (VF0273)
	<i>fliM</i>	flagellar motor switch protein FliM Flagella (VF0051)
		flagellar motor switch protein FliM Flagella (VF0114)
	<i>fliO</i>	flagellar biosynthesis protein Peritrichous flagella (VF0967)
		flagellar biosynthetic protein FliO Flagella (VF0685)
		flagellar protein FliO Flagella (VF0273)
	<i>fliS</i>	flagellar protein FliS Flagella (VF0273)
		flagellar export chaperone FliS Peritrichous flagella (VF0732)
	<i>icsA/virG</i>	autotransporter, actin tail assembly protein IcsA/VirG IcsA (VirG) (VF0121)
	<i>lafX</i>	LafX Lateral flagella (VF0474)
	<i>lfiE</i>	lateral flagellar hook basal body protein, LfiE Lateral flagella (VF0474)
	<i>motA</i>	flagellar motor protein Flagella (VF0051)
	<i>motX</i>	sodium-type flagellar motor component Polar flagella (VF0473)
	<i>motY</i>	probable outer membrane protein precursor Flagella (VF0273)
Nutritional/Metabolic factor (VFC0272)	<i>A1S_0917</i>	transcriptional regulator HemO cluster (VF0636)
	<i>AB57_RS05095</i>	LysR family transcriptional regulator HemO cluster (VF0636)
	<i>ABD1_RS04515</i>	LysR family transcriptional regulator HemO cluster (VF0636)
	<i>ABK1_RS04735</i>	LysR family transcriptional regulator HemO cluster (VF0636)
	<i>ABZ1_RS11145</i>	LysR family transcriptional regulator HemO cluster (VF0636)
	<i>ACICU_RS04565</i>	LysR family transcriptional regulator HemO cluster (VF0636)
	<i>argP</i>	amino acid antiporter ArgP (VF0557)
	<i>ASA_RS16545</i>	ChaN family lipoprotein Heme uptake system (VF0650)
	<i>asbB</i>	petrobactin biosynthesis protein AsbB Petrobactin (VF0584)
	<i>basF</i>	aryl carrier protein BasF Acinetobactin (VF0467)
	<i>bauF</i>	siderophore-interacting protein Acinetobactin (VF0467)
	<i>BJAB0715_RS05225</i>	LysR family transcriptional regulator HemO cluster (VF0636)
	<i>BJE04_RS21780</i>	isochorismate lyase Vibriobactin/Vulnibactin (VF0626)
	<i>entD</i>	enterochelin synthetase component D Ent (VF0562)
	<i>fsIA</i>	lucA/lucC family siderophore biosynthesis protein Rhizoferrin (VF0551)
	<i>fsIC</i>	diaminopimelate decarboxylase Rhizoferrin (VF0551)
	<i>gadC</i>	glutamate:gamma-aminobutyric acid antiporter family protein GadC (VF0555)
	<i>hitA</i>	iron ABC transporter substrate-binding protein HitABC (VF0268)
	<i>hutA</i>	TonB-dependent hemoglobin/transferrin/lactoferrin family receptor Heme receptor HutA (VF0629)
	<i>isdG</i>	iron-regulated surface determinant protein G Isd (VF0015)

	<i>iucA</i>	aerobactin siderophore biosynthesis protein IucD Aerobactin (VF0229)
		aerobactin synthesis protein IucA Aerobactin (VF0123)
		aerobactin Synthetase IucA Aerobactin (VF0565)
	<i>lbtB</i>	LbtB inner membrane protein required for export of Legiobactin Legiobactin (VF0291)
	<i>mbtD</i>	Polyketide synthetase MbtD (polyketide synthase) Mycobactin (VF0299)
	<i>mbtM</i>	Probable fatty acyl-AMP ligase MbtM Mycobactin (VF0299)
	<i>mce1D</i>	virulence factor Mce family protein Mce1 (VF0824)
	<i>PA2384</i>	hypothetical protein Pyoverdine (VF0094)
	<i>PMI_RS01125</i>	IucA/IucC family siderophore biosynthesis protein Proteobactin (VF1252)
	<i>pvcD</i>	paerucumarin biosynthesis protein PvcD Pyoverdine (VF0094)
	<i>VP_RS23105</i>	MFS transporter Vibrioferrin (VF0631)
	<i>ybtT</i>	yersiniabactin biosynthesis thioesterase YbtT Yersiniabactin siderophore (VF1115)
		type II thioesterase YbtT Yersiniabactin (VF0136)
		type II thioesterase YbtT Ybt (VF0564)
Regulation (VFC0301)	<i>agrC</i>	histidine kinase AgrA/AgrC (VF0793)
		sensor histidine kinase AgrA/AgrC (VF0793)
		transcriptional regulator MosR (VF0856)
	<i>mosR</i>	hypothetical protein MosR (VF0856)
		sensor histidine kinase VirR/VirS two component system (VF0729)
Stress survival (VFC0282)	<i>hspX</i>	Heat shock protein HspX (alpha-crystallin homolog HspX (VF0301)
	<i>ureF</i>	urease accessory protein (ureF) Urease (VF0050)
	<i>ureI</i>	acid-activated urea channel protein Urease (VF0050)
Others (VFC0346)	<i>LLO_RS04380</i>	L,D-transpeptidase Enh loci (VF0163)
	<i>LPC_RS04980</i>	
	<i>LPC_RS07250</i>	
	<i>LPG_RS04520</i>	
	<i>LPG_RS06945</i>	
	<i>LPL_RS04745</i>	
	<i>LPP_RS06735</i>	
	<i>mce8A</i>	MCE family protein Mce8 (VF0831)
	<i>ompA</i>	OmpA family protein OmpA (VF0548)
	<i>vapA3</i>	HigA family addiction module antidote protein Virulence-associated proteins (VF0679)